

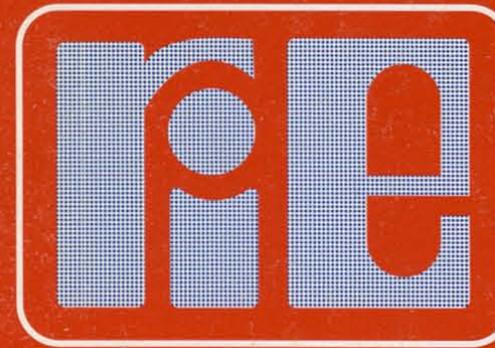


ISSN: 0212-4068
Depósito Legal: B-10235/83

ASOCIACIÓN INTERUNIVERSITARIA
DE INVESTIGACIÓN PEDAGÓGICA
EXPERIMENTAL

REVISTA DE INVESTIGACIÓN EDUCATIVA

N.º 23, 1.º semestre 1994



REVISTA
INVESTIGACIÓN
EDUCATIVA

N.º 23, 1.º semestre 1994



NORMAS PARA LAS COLABORACIONES

El objetivo de la Revista de Investigación Educativa es promover el intercambio de información acerca de investigaciones empíricas de carácter educativo. Todo profesional que desee colaborar en la Revista deberá atenerse a las siguientes indicaciones:

1. Los trabajos deberán ser originales y versar sobre investigación educativa.
2. El autor deberá enviar un original y tres copias mecanografiadas a la redacción de la Revista.
3. La extensión máxima de los trabajos no deberá exceder las 24 páginas en DIN-A4, a doble espacio y numeradas. Se acompañará un abstract de 100 a 175 palabras en inglés y español.
4. Con el fin de simplificar el proceso de confección de la Revista y, sobre todo, de reducir al máximo las erratas y costos, se ruega a los autores enviar el trabajo, además de las copias en papel, en disquette compatible. Se aceptan los siguientes programas de tratamiento de textos: Word Star, Word Perfect, MSWord y otros.
5. Los trabajos recibidos serán sometidos a informe del Consejo Asesor de la Revista (quien decidirá sobre su publicación o no).
6. Se notificará a sus autores los trabajos aceptados para su publicación.
7. Para la redacción de los trabajos se recomienda a los autores que sigan las normas de la A.P.A. (American Psychology Association). (Ver adaptación publicada en el N.º 19, 1.º semestre 1992 de R.I.E.).
8. Las fichas-resumen se cumplimentarán de acuerdo con el modelo propuesto e impreso por la Revista, a cuya sede podrá solicitar los ejemplares.
9. Los números se cierran los días 1 de mayo y 1 de noviembre de cada año.
10. La R.I.E. es de carácter semestral.

REDACCIÓN: REVISTA DE INVESTIGACIÓN EDUCATIVA

Dpto. de Métodos de Investigación y Diagnóstico en Educación
Facultad de Educación

C/. Baldiri i Reixach, s/n. Bloq. D - Piso 3.º
08028 BARCELONA (Spain)

A.I.D.I.P.E.

Asociación Interuniversitaria de Investigación Pedagógica Experimental, creada en 1987

FINES DE LA ASOCIACIÓN

- a) Fomentar, estimular e impulsar la investigación educativa de carácter empírico en todas sus vertientes y ámbitos de aplicación.
- b) Difundir los resultados de esta investigación.
- c) Agrupar a todos los que en España cultivan esta área de conocimiento, como docentes o investigadores.
- d) Colaborar con las autoridades educativas y de política científica en la elaboración de instrumentos de medida, evaluación y de diseños de investigación tendentes a la mejora del sistema educativo y de desarrollo de los mismos.
- e) Favorecer la formación y el perfeccionamiento especializado del personal investigador y de todos aquellos profesionales que intervienen en la investigación educativa.
- f) Velar por la calidad científica y profesional de las actividades relacionadas con la investigación educativa.
- g) Velar por el respeto y cumplimiento de las normas deontológicas de la investigación científica referida al campo psicopedagógico.

DERECHOS DE LOS SOCIOS

Los socios numerarios y fundadores tendrán los siguientes derechos:

- a) Tomar parte en cuantas actividades organice o patrocine la Asociación en cumplimiento de sus fines.
- b) Disfrutar de todas las ventajas y beneficios que la Asociación pueda obtener.
- c) Participar en las Asambleas con voz y voto.
- d) Ser electores y elegibles para los cargos directivos.
- e) Recibir información sobre los acuerdos adoptados por los órganos directivos, sobre las actividades técnicas y científicas que la Asociación desarrolle.
- f) Hacer sugerencias a los miembros de la Junta Directiva de la Asociación.
- g) Impugnar los acuerdos y actuaciones de la Asociación que sean contrarios a los Estatutos, dentro del plazo de cuarenta días y en la forma prevista por las leyes.

Cada socio tiene derecho a recibir la Revista de Investigación Educativa de forma gratuita.

SEMINARIOS

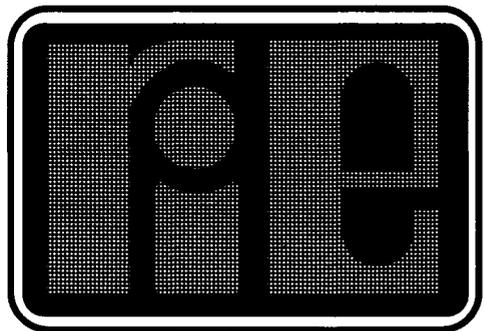
AIDIPE organiza cada 2 años, un Seminario Nacional con un tema central en torno al cual giran las ponencias, comunicaciones, mesas redondas. Los primeros Congresos han sido los de Barcelona (1981), Sitges (1983), Gijón (1985), Santiago de Compostela (1988), Murcia (1990).

PUBLICACIONES

AIDIPE publica las Actas de sus Seminarios, así como La Revista de Investigación Educativa (RIE) con carácter semestral.

Para más información, ver hoja de inscripción adjunta.

REVISTA
INVESTIGACIÓN
EDUCATIVA



VI SEMINARIO DE MODELOS DE INVESTIGACIÓN EDUCATIVA

INVESTIGACIÓN SOBRE DIFERENCIACIÓN EDUCATIVA: LA DIMENSIÓN METODOLÓGICA

JUNTA DIRECTIVA DE A.I.D.I.P.E.

PRESIDENTE: Arturo de la Orden Hoz
SECRETARIA: Julia V. Espín López
VOCALES: Flor Cabrera
Fuensanta Hernández Pina
Carmen Jiménez
Alfredo Pérez Boullosa
Delio del Rincón
M^a Luisa Rodríguez Moreno
Javier Tejedor Tejedor

COMITÉ ORGANIZADOR

Arturo de la Orden Hoz	José Luis Gaviria Soto
Narciso García Nieto	Marta M. González Hernández
Inmaculada Asensio Muñoz	Eloísa López Franco
José Crespo Vasco	José Mafokozi
M ^a José Fernández Díaz	Pedro Municio Fernández
Aurora Fuentes Vicente	Laura Oliveros Martín-Varés
Mercedes García García	Miguel Serra Pérez
José Manuel García Ramos	José Javier Sierra Rubio

COMITÉ CIENTÍFICO

M^a Teresa Díaz Allué
Rafael Carballo Santaolalla
Ángel Lázaro Martínez
Eduardo López López

ENTIDADES COLABORADORAS

Universidad Complutense de Madrid (VII-Centenario)
Facultad de Educación
(C.I.D.E.) Centro de Investigación y Documentación Educativa
Ministerio de Educación y Ciencia (Dirección General de Investigación Científica y Técnica)
Excmo. Ayuntamiento de Madrid

Número: 23
1^{er} semestre, 1994
EDITA

Asociación Interuniversitaria de
Investigación Pedagógica
Experimental (A.I.D.I.P.E.)

Director: **Javier Tejedor**

Directora ejecutiva:
Flor Cabrera

CONSEJO ASESOR:

Margarita Bartolomé
Nuria Borrell
Leonor Buendía
Iñaki Dendaluze
Lisardo Doval
Narciso García
Fuensanta Hernández
Jesús Jornet

Mario de Miguel
Arturo de la Orden
Ramón Pérez Juste
Antonio Rodríguez Diéguez
Francisco J. Tejedor
Carmen Vidal

CONSEJO DE REDACCIÓN:

Margarita Bartolomé
Rafael Bisquerra
Flor Cabrera
Inmaculada Dorio
Julia V. Espín
Pilar Figuera
Ángel Forner
Javier Gil
Fuensanta Hernández
M^a Luisa Rodríguez
Mercedes Rodríguez
Delio del Rincón
M^a Paz Sandín
Antonio Sans

DISTRIBUCIÓN:

Área M.I.D.E.
Facultad de Educación
Campus Espinardo
Universidad de Murcia 30007
Tels. (968) 36 40 67

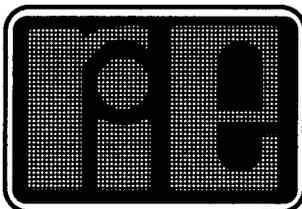
SUSCRIPCIÓN E

INTERCAMBIO CIENTÍFICO:

Dpto. MIDE
Facultad de Educación
Baldri Reixach, s/n. Blq. D-3.^o
08028 BARCELONA

DISEÑO Y MAQUETACIÓN:

Compobell, S.L.
MURCIA



N.º 23

REVISTA
INVESTIGACIÓN
EDUCATIVA

SUMARIO

EDITORIAL 11

PONENCIAS

Ponencia I. Modelos de investigación en la inter-
vención educativa diferencial
por M. Bartolomé, F. Cabrera, J. V. Espín,
M^a A. Marín, D. del Rincón y M. Rodríguez 15

Ponencia II. Perspectivas metodológicas actuales
de la evaluación de programas en el ámbito educa-
tivo
por F. J. Tejedor, A. García-Valcárcel y M^a José
Rodríguez Conde 93

Ponencia III. Modelos de construcción y valida-
ción de instrumentos diagnósticos
por A. de la Orden, J. L. Gaviria, A. Fuentes y
A. Lázaro 129

Ponencia IV. Análisis de datos cualitativos en la
investigación sobre la diferenciación educativa
por E. García, J. Gil y G. Rodríguez 179

COMUNICACIONES A LA PONENCIA I

Evaluando la educación infantil: revisión del mo-
delo y metodología
por Mercedes García García 217

SUMARIO (continuación)

La metodología microgenética: una alternativa metodológica en investigación psico-pedagógica básica por Mercè Garcia-Milà	222
Intervención educativa a través de la investigación cooperativa por Leonor Buendía Eisman y Honorio Salmerón Pérez	226
La investigación-acción como modelo de intervención para el desarrollo de la auto-gestión y la autonomía personal por I. López, S. Romero, T. Padilla, M. Vargas y D. Martín	232
La formación de madres desde la investigación-acción por Teresa Padilla Carmona y M ^a Dolores Martín Benítez	236
Dimensiones implicadas en la inserción laboral: análisis a partir de una escala por Mercè Jariot García y Màrius Martínez i Muñoz	240
Educación ambiental y evaluación en educación primaria por Juliana M. Correa Manfredi y M ^a Teresa Paniagua Aguilar	245
Efectos diferenciales de programas de educación infantil en función de su grado de adecuación a los principios organizativo/didácticos establecidos en el diseño curricular base por M ^a Teresa Aguado Odina	250
Intervención estructural en un caso de retraso escolar y déficit en comportamiento adaptativo por Amparo Martínez Sánchez	255
El sexo y la clase social como determinantes de los intereses profesionales en el último curso de secundaria por José Luis Gaviria Soto	262
Aproximación psicofisiológica al estudio del proceso atencional en niños con síndrome de down por Francisco A. García Sánchez y José M. Martínez Selva	274
COMUNICACIONES A LA PONENCIA II	
Criterios y procedimientos para la selección de indicadores de rendimiento en evaluación institucional universitaria por J. M. Osoro Sierra y L. Salvador Blanco	279

SUMARIO (continuación)

Un estudio sobre indicadores estructurales y calidad percibida en cursos de postgrado por A. Pérez Carbonell, J. M. Jornet, J. González Such y M. R. Ferrández	283
Factores intervinientes en la valoración del profesor por parte del estudiante por J. González Such, J. M. Jornet, A. Pérez Carbonell y M. R. Ferrández	292
Sesgos en la evaluación de la investigación pedagógica y psicológica: una sucinta revisión integrativa por Antonio Fernández Cano	301
Investigación asistida por ordenador: enfoque metodológico por Guillem Bou Bauza y Montserrat Clotet Masana	305
Relaciones interpersonales en centros educativos por Juana M ^a Cruz Martínez y Mercedes Martín Berrido	310
Últimas aportaciones en la evaluación del programa de orientación metacognitiva de la comprensión lectora (Dra. E. Repetto y Col.) por Elvira Repetto, Narciso Barrero, Roberto Castro, Juan Antonio Gil y Concha Martín	314
Evaluación procesual del programa Helios. Un estudio cualitativo de caso múltiple por Samuel Fernández Fernández, Ángel Lázaro Martínez, Pedro Municio Fernández e Isabel Torroba Arroyo	324
Estudio de las actitudes del profesorado hacia la integración de los «escolares diferentes» y su relación con las intenciones de conducta a partir del modelo de Ajzen y Fishbein por Luis Antonio Feliciano García	328
Creencias conductuales del profesorado de EGB sobre los «alumnos diferentes»: un estudio aplicado a la integración del niño artrítico reumatoide desde la perspectiva de la teoría de la acción razonada por Luis Feliciano García y Concepción Riera Quintana	332
Evaluación de un programa de hábitos y técnicas de trabajo intelectual en la transición de la EGB al BUP por Asunción Ayuso Velar, Andrés Luis Haro, Rosario López Barrionuevo, Nieves Oreña Castillo, Francisca José Serrano Pastor	336
Intervención psicopedagógica familiar: adaptación, aplicación y evaluación del programa S.T.E.P./Teen de Dinkmeyer y McKay por I. Bartau y J. M. Maganto	344

SUMARIO (continuación)

Evaluación del practicum de «orientación educativa», a partir de las opiniones de los estudiantes y de los centros de formación por Alfredo Pérez Boullosa, Ignacio Alfaro Rocher, Miguel Ángel Barberá, Laura Ramírez.....	352
Evaluación de los efectos del programa de orientación vocacional y profesional «mirando hacia el futuro» por M ^a del Mar Elías Caraballo y M ^a de los Ángeles Lorenzo Miralles.....	359
COMUNICACIONES PONENCIA III	
Ergo: Características y aplicaciones por I. Inmaculada Asensio Muñoz.....	365
Aplicación de los nuevos enfoques de teoría de la medida en la construcción de una prueba diagnóstica de comprensión verbal por F. Javier Murillo Torrecilla.....	368
Modelos Logit en la validación y el análisis de registros de observación por Juan Carlos Tójar Hurtado.....	371
T.O.P.S.I.S.: Un método de decisión multiatributo como técnica evaluativa por Rafael Galante Guille.....	375
Análisis del funcionamiento diferencial de ítems de respuesta graduada: Una aproximación metodológica por F. Javier Murillo Torrecilla y María Castro Morera.....	379
Procedimiento para la determinación arbitraria de los valores de una matriz de correlaciones por José Luis Gaviria Soto.....	384
Interpretación de la fiabilidad en el análisis documental mediante elecciones dicotómicas y su incidencia en la investigación cualitativa por Santiago Nieto Martín.....	391
La utilización de los procedimientos de comparaciones múltiples en la investigación educativa en España por A. Sáez, J. M. Suárez, F. Aliaga y R. M. Bo.....	396
Construcción de pruebas criterioles de aula por Leonor Buendía Eisman y Honorio Salmerón Pérez.....	405

SUMARIO (continuación)

Proceso de elaboración de un instrumento diagnóstico de la experiencia docente. (Estudio exploratorio) por Concepción Riera Quintana	411
Diseño de un instrumento de análisis de la enseñanza comunicativa de lenguas extranjeras por Mercedes Verdú Jordá y Enrique Iglesias Verdegay	415
Evaluación dinámica de los procesos y estrategias cognitivas en niños deficientes mentales por Santiago Molina García y Ana Arráiz Pérez	423
Proceso de construcción y validación del TEF por Cristina Cardona Moltó, M ^a Ángeles Martínez Ruiz y Narciso Sauleda Parés ..	428
Diagnóstico de las diferencias étnicas y de los procesos desarrollados en la educa- ción primaria. Percepción del profesorado Por Flor Cabrera, Julia V. Espín, M ^a Ángeles Marín y Mercedes Rodríguez Lajo ..	433
Diagnóstico del contexto familiar de alumnos de minorías étnicas por Flor Cabrera, Julia V. Espín, M ^a Ángeles Marín y Mercedes Rodríguez Lajo ..	440
Escala de actitudes del profesorado ante la educación multicultural por Flor Cabrera, Julia V. Espín, M ^a Ángeles Marín y Mercedes Rodríguez Lajo ..	445
El mapa cognitivo de las profesiones por M ^a Teresa Muñoz Sastre	451
Estimación del vocabulario total conocido por un escolar por A. Suárez y P. Meara	455
Análisis del test PMS de Raven en los escolares de Galicia por Eduardo Abalde Paz y M. Miguel Muñoz Cantero	458
Un primer acercamiento a las estrategias de resolución en la tarea de cubos por N. Orellana, J. M. Suárez, R. Bo y F. Aliaga	461
Aproximación al mapa de dificultad del laberinto 6 de la escala WISC-R por R. M. Bo, J. M. Suárez, N. Orellana y A. Sáez	467
Algunos datos sobre la estructura de dificultad del elemento caballo de la subprueba rompecabezas del WISC-R por L. Salavert, J. M. Suárez, C. Belloch y N. Orellana	474

SUMARIO (continuación)

Un cuestionario para identificar niños altamente creativos por Judith Serrano Segura	482
Escala de creencias valorativas de los padres por Trinidad Donoso	485
Un instrumento para evaluar la conducta exploratoria en el desarrollo de la carrera profesional por T. Donoso, P. Figuera y M. L. Rodríguez	490
Alternativa al «Children's depression inventory» (CDI): Procedimiento y metodología por L. Cabrera Pérez y P. García-Medina	497
COMUNICACIONES A LA PONENCIA III	
La teoría de la información en el análisis de datos cualitativos por Xabier de Salvador González	501
Posibilidades de WordPerfect 5.1. para el análisis de datos cualitativos por Luis Carro San cristóbal	505
Análisis estadístico de datos cualitativos textuales: El enfoque lexicométrico por Javier Gil, Eduardo García, Gregorio Rodríguez y Ana Corrales	510
Formulación y comprobación de hipótesis cualitativas con ayuda del programa «AQUAD» por Javier Gil, Eduardo García, Gregorio Rodríguez y Ana Corrales	515
NUDIST: Una herramienta informática para el análisis de datos cualitativos por Ana Corrales, Gregorio Rodríguez, Javier Gil y Eduardo García	522
NUDIST versus AQUAD: Dos posibilidades informáticas para el análisis de datos cualitativos por Gregorio Rodríguez, Javier Gil, Ana Corrales y Eduardo García	529
Categorización y complejidad en la investigación cualitativa por Xavier Gil Quesada	535
Problemas técnicos enunciados en el ámbito de la investigación etnográfica sobre la diferenciación educativa por Fernando Sabirón Sierra	538

SUMARIO (continuación)

La construcción del conocimiento en los estudiantes universitarios
por Concepción Medrano Samaniego 543

El estudio del autoconcepto en chicos y chicas con técnicas cualitativas
por M^a Ángeles de la Caba Collado 547

Estudio de necesidades para una adaptación curricular
por M^a Celia Isorna Sánchez, Rosa Morillo Peña, M^a José Vecino Jiménez 551

SIMPOSIUMS

Simposium I. ¿Orientación por programas o por servicios?
por M^a Teresa Díaz Allué 557

La intervención por programas: ¿Una simple innovación tecnológica?
por Víctor Álvarez Rojo 557

Y, ¿los servicios sirven?
por Benito Echeverría Samanes 565

Orientación por programas: Un caso práctico
por Josefina Hernández Fernández 570

Simposium II. ¿Por qué diferenciar los términos *evaluación* y *diagnóstico*?
por P. Muncio 577

Diagnóstico y evaluación
por Ignacio J. Alfaro Rocher 577

¿Diagnóstico versus evaluación?
por Eduardo García Jiménez 584

¿Por qué diferenciar los términos *evaluación* y *diagnóstico*?
por Ramón Pérez Juste 588

¿Se evalúa y/o se diagnostica?
por Ángel Lázaro Martínez 592

COMUNICACIONES AL SIMPOSIUM I

Diseño y planificación de un programa de intervención en orientación universitaria
por Francisca Castellano Moreno y José Antonio Delgado Sánchez 595

Pautas para la evaluación de un programa de orientación
por J. Hernández Fernández y P. Martínez Clares 598

SUMARIO (continuación)

La satisfacción/insatisfacción en el trabajo de los orientadores escolares: Implicaciones profesionales por Luis Sobrado Fernández y Ana M ^a Porto Castro	605
Valoración del perfil formativo-profesional y del acceso del logopeda de los equipos psicopedagógicos de sector escolar por Luis Sobrado Fernández, Ana M ^a Porto Castro y Leonor Filgueira Domínguez	610
La labor orientadora del profesor universitario desde la perspectiva del alumnado y del profesorado por Feli Arbizu Bakaikoa	614
COMUNICACIONES AL SIMPOSIUM II	
¿Diagnóstico pedagógico o evaluación educativa? Hacia una clarificación conceptual y terminológica por Raquel-Amaya Martínez González	623
Estudio longitudinal del temperamento infantil de 0 a 12 meses: Resultados preliminares de una aproximación metodológica interdisciplinar por Francisco A. García-Sánchez, Julio Pérez-López y José A. Carranza	631
Una experiencia de diagnóstico de los procesos de enseñanza-aprendizaje en la Universidad a través de la investigación-acción por S. Romero Rodríguez, D. Ballester Angulo, D. Flores García, J. García Murillo, L. Giráldez Sánchez, M ^a D. Macipe Fernández, E. Rubio Rodríguez y A. Sánchez Sánchez	636
Diagnóstico-evaluación: Cooperación y complementariedad por Santiago Castillo Arredondo	640
La evaluación de la lectura: Variables implicadas en el reconocimiento de palabras por Sylvia Defior, Fernando Justicia y Francisco J. Martos	643
La figura de las letras cursivas por Ignacio Alfaro, Francisco Secadas	647
Diferenciación del riesgo en los conductores. Estudio de perfiles por Miquel Amador Guillen	654
Índices cardiovasculares durante pruebas de evaluación en el aula y en el laboratorio por P. A. Caballero García, N. Navarro Adelantado y F. A. García Sánchez.....	660
Ser de Ciencias..., ser de Letras por Tomás Escudero Escorza	665

EDITORIAL

Con un retraso mayor del que hubiéramos deseado, debido sobretudo a la espera de las prometidas ayudas de financiación, publicamos en este número los trabajos presentados al VI Seminario de AIDIPE dedicado al tema «Investigación sobre diferenciación educativa: La dimensión metodológica», celebrado en Madrid y organizado por nuestros compañeros del Departamento MIDE de la Universidad Complutense. Como en ocasiones anteriores, quiero agradecer en nombre de todos los asistentes y especialmente en nombre de los socios de AIDIPE, los generosos esfuerzos de nuestros compañeros que hicieron posible una vez más este fructífero encuentro, posibilitándonos la necesaria puesta al día en algunos temas y el intercambio de opiniones sobre los temas tratados.

Aunque han sido varias las causas que han motivado el retraso en la publicación, ha sido sin duda nuestra propia ambición por conseguir financiación para la totalidad de los trabajos presentados al Seminario (ponencias y comunicaciones) uno de los motivos determinantes de tal situación. Al fin, la perseverancia ha dado sus frutos y podemos, sobrepasando los compromisos adquiridos en el período informativo previo a la realización del Seminario, publicar la totalidad de los trabajos presentados.

El tema elegido para tratar en el Seminario era, sin duda, intencionalmente amplio, a fin de dar cabida a las diferentes concepciones y tratamientos de la muy diversa problemática de la realidad educativa, aunque se tratase exclusivamente desde la óptica de los procesos de intervención, en su doble conceptualización: sustantiva o de contenido diferenciado y metodológica.

La estructura de los temas tratados en el Seminario se concretó en cuatro ponencias, a las que se vincularon numerosas comunicaciones, y dos symposium o foros de debate. Las cuatro ponencias «marco» fueron las siguientes:

I. Modelos de investigación en la intervención educativa diferencial, elaborada por un conjunto de profesores del Área MIDE de la Universidad de Barcelona. Se aceptó la presentación de 11 comunicaciones que fueron presentadas entonces y que ahora publicamos.

II. Perspectivas metodológicas actuales de la evaluación de programas en el ámbito educativo, elaborada por un grupo de profesores de la Universidad de Salamanca. Se aceptaron 14 comunicaciones vinculadas a esta ponencia.

III. Modelos de construcción y validación de instrumentos diagnósticos, elaborada por profesores del área MIDE de la Universidad Complutense. Se aceptaron 26 comunicaciones vinculadas a esta ponencia.

IV. Análisis de datos cualitativos en la investigación sobre diferenciación educativa, elaborada por los compañeros de MIDE de la Universidad de Sevilla. Se seleccionaron 11 comunicaciones.

En los dos symposium celebrados se trataron los temas ¿Orientación por programas o servicios? y ¿Por qué diferenciar los términos Evaluación y Diagnóstico? Un total de 13 comunicaciones se aceptaron para comentar en estos foros de debate.

Confiamos en que el material presentado sea de interés para la comunidad educativa y, especialmente, para los lectores de la Revista. En todos los casos, los autores aceptan correspondencia sobre las opiniones expuestas.

Francisco Javier Tejedor
Presidente AIDIPE

PONENCIAS

PONENCIA I

MODELOS DE INVESTIGACIÓN EN LA INTERVENCIÓN EDUCATIVA DIFERENCIAL

por

Margarita Bartolomé, Flor Cabrera, Julia Victoria Espín, María Ángeles Marín,
Delio del Rincón y Mercedes Rodríguez Lajo
Universidad de Barcelona

INTRODUCCIÓN

Intentar acotar el extenso campo de esta ponencia no resulta tarea fácil. ¿Qué entendemos por «intervención educativa diferencial»? Hace una década planteábamos la misma pregunta. «¿Aludimos a que existen tipos de educación distintos, situaciones educativas diversas, sujetos que reciben o imparten la educación, con características diferentes? ¿Son las variables de entrada, proceso o producto las responsables del vocablo diferencial?» (Bartolomé, 1983, 5). Ya entonces se indicaba la necesidad de responder a estos interrogantes desde la consideración de la relación educativa como un proceso dinámico e interactivo, donde las diferencias iniciales condicionan y al tiempo son susceptibles de modificación en el curso de la intervención educativa. Ésta a su vez, desde el *principio de diferenciación* y desde un *enfoque constructivista*, debería flexibilizarse e individualizarse con proyectos curriculares adecuados a los conocimientos previos, características peculiares de las personas y necesidad de ayuda requerida.

Algunas de las reflexiones hechas sobre la conceptualización de la Pedagogía Diferencial y su vertiente metodológica, en las Primeras Jornadas de esta materia (Madrid, 11 y 12 de marzo de 1988), inciden de manera especial en este punto (Tejedor, 1988, 591-193).

Sin embargo, la investigación en este campo no puede restringirse a la evaluación sistemática de los modelos de intervención que vayan surgiendo y que permitan una educación diferenciada sino a *investigar las respuestas reales que se están produciendo en la intervención educativa, analizando en qué medida se adecúan dinámicamente a las diferencias individuales y grupales y están o no produciendo efectos diferenciadores.*

La acotación del tema se ha hecho pues a partir de la consideración central de la intervención educativa diferencial desde esa perspectiva amplia.

Tres han sido las áreas comprendidas en nuestra aportación. Dos de ellas hacen referencia a diferencias grupales que por su relevancia y actualidad como objetos de estudio en el ámbito de la investigación y actuación educativas nos ha parecido conveniente seleccionar: *las diferencias de género así como las étnicas y socioculturales*. Sin identificar en absoluto estas dos últimas clasificaciones, el hecho de incluir en un mismo apartado las diferencias étnicas y las socio culturales se justifica desde la urgencia y preocupación que despierta en el horizonte europeo la atención a las minorías étnicas, procedentes en muchos casos de la migración exterior, que sufren a su vez, una marginación social y cultural. Baste recorrer las bolsas de población marginada de las grandes ciudades europeas y descubrir el porcentaje de población inmigrante que las compone. Esta asociación es tan estrecha, que algunos antropólogos han acuñado el término de diferencias *socioétnicas* para referirse a las que se producen al combinar el estatus social y la etnia. La minorías étnicas económicamente potentes no viven la misma problemática social, cultural y educativa que estos grupos marginados. De todos es conocido que los árabes de Marbella tienen muy poco en común con los últimos magrebís que atravesaron en una yola el estrecho. Ya en 1972 Marshall hacía referencia a la asignación en Brasil a una raza u otra de las personas procedentes de cruces raciales, dependiendo del estatus social que ocupaban.

Muy diferentes son las cuestiones educativas suscitadas por las denominadas minorías históricas, la afirmación de cuya identidad puede asociarse a los procesos desencadenados por las políticas públicas, para cuyo análisis e investigación se sugieren algunos elementos en uno de los apartados de esta ponencia.

Si bien la cercanía de recientes congresos a nivel nacional sobre educación multicultural e intercultural podrían haber desaconsejado la inclusión de este tema, la necesidad de hacerlo avanzar y el hecho de que nuestro Departamento se halle embarcado en el mismo a través de una amplia investigación, nos han animado a incluirlo, aunque con una restricción: no hacemos referencia explícita a las numerosas investigaciones sobre bi y multilingüismo que corrientemente se asocian a una educación multicultural. Un análisis detallado de los modelos de educación multilingüe (Marín, 1992) y de educación multicultural (Selby, 1992) desarrollados en los procesos educativos de distintos países de Europa, en las últimas décadas, (Siguán, 1992), permite constatar la no coincidencia en la evolución de los modelos. Creemos que desde otros Departamentos del Área, que tienen ya una larga tradición en el estudio diferencial del Lenguaje, (como el grupo del Área MIDE de la Universidad del País Vasco) o de otras instituciones, como el ICE de la Universidad de Barcelona, podrán ir enriqueciéndose y ampliándose las aportaciones presentadas en el X Congreso Nacional de Pedagogía (1992).

También las diferencias educativas asociadas a la edad han quedado fuera de este trabajo, en razón de su extensión. Indirectamente se aludirá a esta variable al

aludir a la evolución de las diferencias de género, étnicas o socioculturales en los distintos niveles educativos.

En el campo de las *diferencias individuales*, nos hemos ceñido fundamentalmente a las aplicaciones al medio escolar de los constructos procedentes de la psicología cognitiva, dada su incidencia en los planteamientos actuales de la Reforma educativa en España. No se incluyó, por considerarlo inicialmente parte de otra ponencia, las investigaciones sobre adaptaciones curriculares para sujetos con necesidades educativas especiales.

Hemos tenido en cuenta además, la posibilidad de investigar la intervención educativa diferencial, en los campos anotados anteriormente, a *distintos niveles*: macro, meso y micro. Nos referiremos, tanto a los planteamientos llevados a cabo por las *políticas públicas* y tendentes a abordar educativamente el tema de las diferencias, para evitar que éstas se conviertan de hecho en desigualdad de oportunidades, así como a investigaciones desarrolladas en el ámbito de las *instituciones educativas o del aula*.

Pasemos pues a presentar los cuatro apartados de que constará esta ponencia.

1. INVESTIGACIÓN Y POLÍTICAS PÚBLICAS QUE PROMUEVEN UNA INTERVENCIÓN EDUCATIVA DIFERENCIADA

1.1. La investigación de las políticas públicas

La relación entre investigación y políticas educativas tiene una historia reciente. Será a partir de los años 60 cuando la evaluación de los programas y de las políticas comienza a ser un campo de estudio esencial y de rápida extensión en EE.UU. y, posteriormente, en los años 70 y siguientes, en los demás países del mundo desarrollado. Paralelamente a las preocupaciones de naturaleza directamente educativa, se inicia en el ámbito de la administración pública en los EE.UU., un movimiento de estudio y reflexión por buscar la eficacia y productividad de las decisiones de los gobiernos y actuaciones públicas. Se desarrollan nuevas tecnologías sobre el quehacer público que sentarán las bases teóricas de las que ha emergido la contemporánea teoría de la administración pública. Dentro de las tendencias dominantes en este campo hoy, destaca el «análisis de políticas públicas». —*policy analysis*— (Subirach, 1989; Ballart 1992), como una técnica instrumental útil para la elaboración y formulación racional de las políticas públicas. El gran desarrollo del *policy analysis* durante estos últimos 20 años, dará lugar a que la investigación educativa, bajo la forma esencial de *investigación evaluativa*, pase a jugar un papel esencial en el terreno de las decisiones políticas (Nagel, 1990).

Los ámbitos de estudio que han ocupado los esfuerzos de los investigadores en relación con las políticas y actuaciones públicas se concretan en trabajos referentes a: a) investigación evaluativa sobre políticas y la utilidad de los resultados de la evaluación en las decisiones políticas; b) estudios sobre indicadores de gestión y evaluación de políticas; y c) estudios de costes.

1.2. Ámbitos de interés de la investigación evaluativa relacionada con políticas públicas sobre educación diferencial

Se reconoce al informe Coleman (1966)¹ como el trabajo pionero más influyente en las decisiones políticas del campo educativo, como afirma Langbein (1980) — en Ballart (1992)— el *Coleman Report* proporcionó una credibilidad que no tenían otras evaluaciones anteriores, tanto por los métodos de investigación utilizados como por la forma de aproximarse al problema de la igualdad de oportunidades en educación.

Desde 1960 la investigación ha tenido un impacto significativo sobre las políticas que, en concreto, ajustándonos a nuestra línea de trabajo sobre las intervenciones educativas diferenciales, podemos ubicarlas en los siguientes grandes ámbitos (Grant, 1992; Cibulka, 1992):

- a) Políticas sobre educación multicultural
- b) Políticas antipobreza con una preocupación de incidencia netamente educativa educación compensatoria - o de manera indirecta mediante ayudas o servicios sociales.
- c) Políticas educativas para la igualdad de oportunidades entre los sexos.

A. Educación multicultural

Según McCarthy (1990), las estrategias educativas que se derivan del discurso político en este ámbito obedecen a tres modelos de actuaciones diferentes: a) políticas que potencian la *comprensión cultural*, la idea central es el desarrollo de una mayor sensibilidad entre los profesores y los estudiantes hacia las diferencias étnicas; b) políticas que promueven la competencia cultural, donde se insiste que profesores y estudiantes desarrollen una competencia lingüísticas y cultural relativas a otros grupos étnicos; y c) políticas que promueven la *emancipación cultural*, es decir, el reconocimiento y la inclusión de la cultura de las minorías en los currícula escolares por su reconocida influencia en la autoestima y en los logros académicos de las minorías, con su consecuente repercusión más allá de la vida escolar.

Si bien se reconoce la educación multicultural a nivel de discursos y declaraciones políticas, son *escasos los estudios e investigaciones* que se emprenden a este nivel. Una revisión del sistema ERIC de los informes sobre policy analysis hecha por Crumpton (1992), dio como resultado que sólo 50 informes usaban el término multicultural, y 39 de ellos específicamente el de educación multicultural. Sin embargo de aquellos 39 informes, el autor encontró sólo tres que se referían a la realización de estudios a nivel de estado sobre educación multicultural. Los tres informes eran estudios descriptivos por encuesta a los departamentos de edu-

1 Con frecuencia se le denomina así al título original de la obra de Coleman: *Equality of educational opportunity*.

cación de los estados a fin de conocer sus planteamientos en educación multicultural —Freeman (1983), The American Association of Colleges for Teacher education (1978) y Mitchell B (1985)—. No eran estudios empíricos que investigaran los efectos de determinadas políticas, sino *trabajos descriptivos* de las regulaciones y políticas que se estaban desarrollando en los distintos estados. Lo que llama la atención, sin embargo, es que cuatro años más tarde Mitchell (1988) comprueba en su estudio de seguimiento que apenas había variado la situación. Los trabajos de Mitchell (1984) y Gollnick y otros (1976) sobre el estado de la cuestión en el educacional policy analysis que ofrecen una revisión de los estudios étnicos y de educación multicultural en EE.UU., señalan la *escasez de estudios políticos* en el área de la educación multicultural, y los escasos presupuestos dedicados al apoyo de iniciativas en educación multicultural, a pesar de que median ocho años de diferencia entre ambos estudios.

Crumpton (1992) siguiendo a Gollnick (1976) y a Mitchell (1988) sugiere las cinco áreas de investigación que deberían promoverse desde los departamentos de educación en materia de educación multicultural:

- Identificar y definir standards de educación multicultural que ayuden a formular objetivos de educación multicultural.
- Promover iniciativas de acercamientos a la educación multicultural.
- Investigar y dar a conocer los resultados de las investigaciones sobre programas efectivos en educación multicultural.
- Promover una política de incentivos y sanciones a fin de estimular el desarrollo de programas de educación multicultural.

De interés es la propuesta de Lynch (1986) —recogida por Gollnick (1992)— de distinguir las ideologías y paradigmas subyacentes en los análisis sobre políticas multiculturales. El autor, después de estudiar las políticas de varios estados, distingue tres tipos de ideologías: económica, democrática y de compromiso social en función del tipo de valores, conocimientos, estructura social y controles sociales que subyacen.

B. La lucha contra la pobreza

La educación compensatoria

Posiblemente el campo donde la investigación educativa se ha relacionado más con preocupaciones políticas e interés público ha sido en los estudios evaluativos realizados en EE.UU. sobre programas de educación compensatoria, que obedecían a las políticas de igualdad de oportunidades (The Economic Opportunity Act de 1964) y de lucha contra la pobreza (antipoverty program del presidente Johnson).

Dos estudios nacionales realizados sobre la efectividad del programa Head Star son tomados como ejemplos por Anderson (1984) para ilustrar la *dimensión política de investigación evaluativa* cuando se sitúa en la arena política: el estudio nacional encargado por la División de Evaluación de OEO (Office Economic opportunity) a

The Westinghouse Learning Corporation y a la universidad de Ohio y el estudio realizado por el Department of Health, Education and Welfare.

El estudio de Westinghouse se inició ya con una discusión entre el personal de OEO y el personal del programa Head Start, respecto a los objetivos de la evaluación, al diseño de investigación y a los tipos de medidas a utilizar. Tratándose de una *policy evaluation* más que una evaluación de programa², las directrices de la investigación fueron las marcadas por la división de evaluación de OEO. Se valoró sólo objetivos referentes al desarrollo cognitivo, se utilizó un *diseño experimental* con grupo control con medición de resultados a corto plazo, frente a los intereses del personal del programa porque se valoraran otros objetivos y sobretodo se utilizara un diseño de estudios longitudinal.

Los resultados de la evaluación de Westinghouse donde se ponía de manifiesto la poca efectividad del programa saltaron a la arena pública cuando el presidente Nixon alude a la pobreza de sus resultados —téngase en cuenta que el programa nace bajo la presidencia del demócrata Johnson—. Ello dio lugar a que los contrarios al programa presionaran para que se hicieran públicos los resultados del estudio, y acusaran a la oficina de evaluación de la OEO de querer tergiversarlos buscando nuevos reanálisis. De otra parte, los partidarios del programa sometían a terribles críticas la metodología seguida en el estudio de Westinghouse—diseño, medidas utilizadas, muestreo, estadísticos— para justificar la poca credibilidad que merecían sus resultados negativos. A la vista de esto, podemos cuestionarnos: ¿se hubiesen hecho públicos los resultados de la evaluación negativa si el color político dominante en lugar de republicano hubiese sido el demócrata?, ¿se hubiera criticado tanto la metodología seguida en el estudio y se hubieran hecho tantos reanálisis si los resultados hubieran coincidido con las mejoras esperadas?

Las respuestas a las cuestiones anteriores son fáciles de imaginar a la luz de lo que ocurrió con el otro estudio nacional, realizado por el Department of Health, Education and Welfare diez años más tarde. Utilizándose un *diseño longitudinal* se mostró que el programa tenía efectos significativos a largo plazo para los estudiantes que habían pasado por él: necesitaban menos clases de recuperación, disminuían su repeticiones de curso y aumentaban el número de los que continuaban estudios superiores. A partir de aquel momento, el Head Start fue aclamado como un buen programa, como dice Anderson (1984), resultaba una tarea fácil para todos —políticos, gobierno, grupos de opinión etc.— aceptar unos resultados de investigación que coinciden con el sentimiento de simpatía que profesionales y público tenían del programa. El propio autor, ante la diferencia de resultados en ambos

2 Ballart (1992) distingue entre la evaluación de programas y evaluación de políticas precisamente por el diferente uso que tienen los resultados. Cuando se evalúa políticas se pretende servir a decisiones sobre directrices políticas y se intentaría responder a cuestiones relacionadas con la generalización de una política en cuanto aproximación a un determinado problema. En cambio, en una evaluación de programa se pretende mejorar la actuación en sí misma y sus resultados tienen relevancia a nivel de gestión de programas.

estudios, se pregunta, ¿se trata en realidad de una diferencia en la efectividad del programa o es el resultado de diferencias en las metodologías y técnicas de análisis utilizadas en las investigaciones?

En conexión con la idea anterior, es reconocida la abundante literatura que refiere estudios evaluativos sobre programas de educación compensatoria, análisis y reanálisis de sus efectos, de la misma manera que han proliferado los metaanálisis, y todo ello animado posiblemente por lo poco conclusivos e incluso contradictorios de los resultados obtenidos sobre los potenciales efectos de los programas compensatorios. Trochim (1983) apoyándose en sus propias revisiones y en los resultados de una encuesta nacional sobre las escuelas con programas de compensatoria, identifica tres diseños fundamentales en los estudios evaluativos sobre educación compensatoria: el *modelo de referencia normativa*, el *modelo de grupo control* y el *modelo de la discontinuidad de la regresión*.

El mismo autor después de la revisión de más de 100 estudios evaluativos, observa que existe una tendencia a que los resultados sobre los efectos de los programas sean más altos en el caso de los diseños que siguen el modelo de referencia normativa que los que siguen el de la discontinuidad de la regresión. El autor analiza como estas diferencias en los resultados pueden ser explicadas por los sesgos propios a cada diseño utilizado: los modelos de referencia normativa tienden a sobreestimar los efectos del programa y los de la discontinuidad de la regresión tienden a subestimarlos.

El estudio longitudinal dirigido por Slavenburg (1991) representa un modelo compensivo, para la investigación evaluativa de proyectos de educación compensatoria. Utiliza un cuidadoso y estudiado *diseño causal* en una *investigación longitudinal* de 11 programas de educación compensatoria (6 escolares y 5 familiares) siguiendo una cohorte de alumnos de primaria, desde el primer al octavo grado, en Amsterdam (The Dutch Education and Social Environment Project). Para poder responder no sólo a la cuestión fundamental de si el programa conseguía sus objetivos —aumentar el rendimiento escolar—, sino también determinar las causas de la no eficacia, y determinar la eficacia relativa de los programas, el equipo de evaluación decidió elaborar el modelo causal presentado en la figura 1. Según el modelo, la puntuación obtenida en el grado 8 se explica por las puntuaciones obtenidas en los grados anteriores, por las características cognitivas y socio-emocionales de los alumnos (inteligencia, hábitos de trabajo etc.), por el nivel de implementación de los programas escolares de compensatoria y por la conductas del profesor (tiempo de enseñanza, tipo de enseñanza, etc). Las características cognitivas y emocionales del alumno se explican por las mismas características en cursos anteriores y por las características educativas de las familias (clima pedagógico, nivel de expectativas, etc.). Las características educativas de las familias se explican por las existentes en los años anteriores, por las características estructurales de la familia (profesión y nivel de educación del padre, de la madre, etc.), y por el nivel de implementación de los programas dirigidos a la familia. Se consideraron como variables exógenas la inteligencia genotípica de los alumnos y las ya mencionadas, como estructura de la

familia, implementación del programa y conductas de los profesores; entre las dos primeras existe una intercorrelación. La construcción del modelo fue precedida de cuidadosos estudios que garantizaran tanto la validez de contenido como la empírica de las relaciones hipotetizadas y las que subyacen bajo cada una de las facetas consideradas. A título de ejemplo la figura 2 muestra el análisis causal comprobado previamente (incluso con análisis factoriales a fin de reducir variables) para validar los path entre las variables que se identificaron como definitorias de las «características educativas de la familia».

Programas y acciones sociales

Las políticas antipobreza han dado como resultado, como señala Espín (1991), no sólo programas de naturaleza eminentemente instructiva, sino programas y acciones de ayudas sociales que de forma directa o más indirectamente deberían redundar en una mejora de las posibilidades educativas de los estudiantes con déficits económicos-sociales-culturales (aumento de recursos, dotaciones de becas, sea a nivel de familia, de escuela o de alumno). Dos tipos de programas que se sitúan en esta línea política son el de *Negative Income Tax Program* en EE.UU. —en Maynard y Murnane (1981)— y el desarrollado por el gobierno de Brasil y Banco Mundial, *Northeast Rural Primary Project (EDURURAL)*, los cuales representan diseños clásicos de investigación en el estudio de las políticas públicas. Son ilustrativos de la discusión sobre la eficacia de los estudios a gran o pequeña escala en el terreno de la evaluación de políticas.

El Negative Income Tax Program (NIT) pretende estudiar los efectos en el rendimiento académico de programas políticos que inciden directamente en el medio familiar, en este caso mediante un salario familiar suplementario. El análisis de los efectos del NIT se realizó mediante *diseños experimentales muy controlados*. Se trabajó muy bien el modelo de asignación de las familias al grupo experimental y al grupo control a fin de eliminar los efectos de la variable «asignación», de tal manera que los resultados de los diferentes experimentos resultaron consistentes. El análisis de los datos se hizo mediante la comparación de la regresión entre ambos grupos, tomando como sucesivas variables dependientes los rendimientos en lengua, en matemática, los niveles académicos conseguidos y las faltas de asistencia, y como variables independientes los grados escolares, los años que las familias llevaban en el programa, y el nivel de pobreza de las familias.

Por otra parte, los pobres resultados obtenidos por Utarbigon y Hanushe (1992) en la valoración del programa EDURURAL ponen de manifiesto los problemas derivados de los estudios a gran escala. EL programa EDURURAL consistió en ayudas de recursos materiales a las escuelas de zonas desfavorecidas del norte de Brasil, bajo el supuesto de que tales recursos mejorarían la calidad instructiva de las escuelas. Si bien en un principio se planificó la implementación del programa seleccionándose, totalmente al azar, las escuelas experimentales y las de control en 218 municipios rurales de zonas muy desfavorecidas del Brasil; en la práctica, la

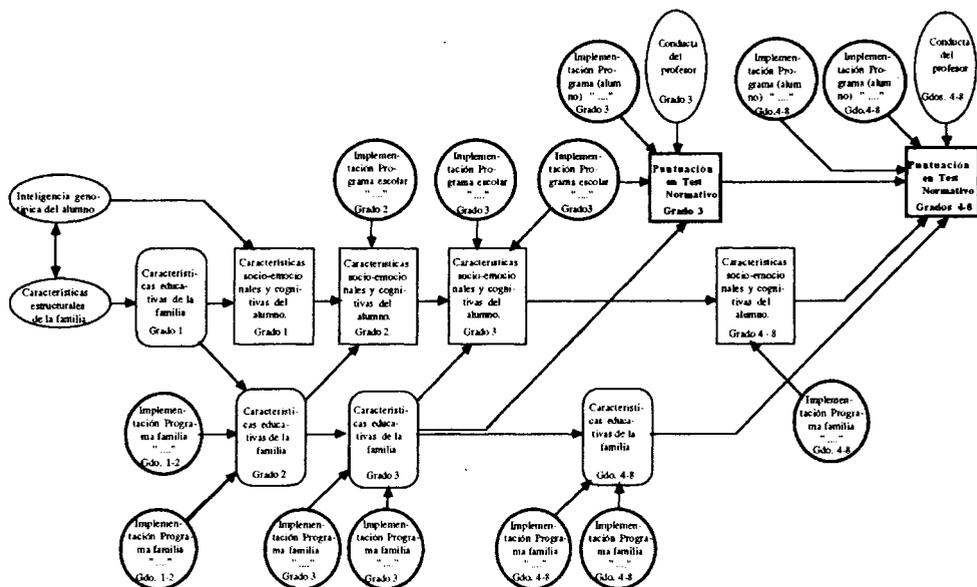


Figura 1. Modelo causal del estudio evaluativo del proyecto de educación Compensatoria de O.S.M. Amsterdam. J. H. Slavenburg (1991).

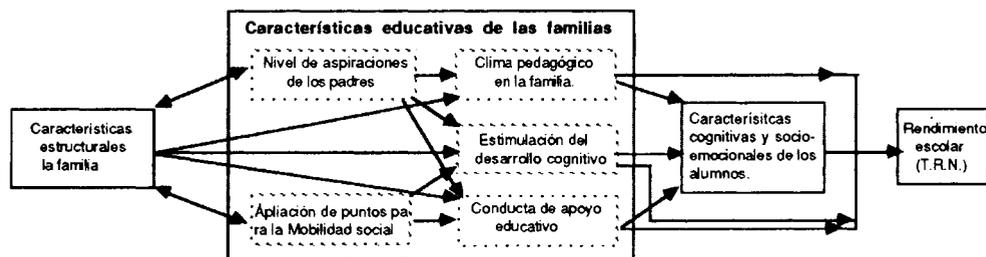


Figura 2

Modelo parcial de las características educativas de la familia (Sluvenburg, 1991).

enorme amplitud geográfica en la que se actuaba impidió controlar la «no contaminación» entre los diferentes tipos de escuelas, así como asegurar una igual implementación del programa. Ante ello, no es de extrañar que las sucesivas valoraciones que se hacían de los efectos del programa (*estudios de seguimientos sobre el rendimiento* de los alumnos de ambas escuelas durante la década de los ochenta) daban resultados diferentes y contradictorios según los Estados. Los autores reconocen que el resultado de sus esfuerzos no fue en la dirección evaluativa, sino en el aumento del conocimiento sobre el fenómeno educativo y, en consecuencia saber lo que se hace en educación y que políticas deberían seguirse.

C. Las políticas de igualdad de oportunidades entre los sexos

Dentro del movimiento de igualdad de oportunidades generado en los años 60 el tema de la igualdad entre los sexos constituía un tópico más a tratar en las políticas educativas, pero en aquellos años el énfasis se colocaba en las desigualdades socioeconómicas de los alumnos y la preocupación era la extensión de la enseñanza como motor de la prosperidad económica. También a medida que se ha extendido la enseñanza, se pone de relieve que las diferencias entre los estudiantes también son mayores — sociales, étnicas, sexuales, intelectuales etc. — No extraña por ello, que cuando en noviembre de 1990 el Comité de Educación de la OCDE reunido a nivel ministerial bajo el tema de una «calidad para todos» destaque como temas prioritarios para las políticas de los diferentes gobiernos de los estados miembros: la igualdad entre los sexos, la integración de los alumnos con déficits, el pluriculturalismo y la enseñanza bilingüe.

Desde mediados de 1970 en la Europa del Mercado Común se han emprendido una serie de acciones concretas a fin de asegurar la igualdad de trato entre hombres y mujeres en materia de educación y acceso al trabajo, incluida la remuneración. En la actualidad se concretan en Planes específicos de actuación en este campo; en estos momentos transcurre el Tercer Programa que cubre el periodo 1991-1995. Los sectores clave donde han incidido las acciones políticas desarrolladas en dichos planes de igualdad son: a) el ámbito legal, con la finalidad de lograr una mejor aplicación, implementación y desarrollo de una legislación no discriminadora de la mujer; b) el campo de la educación, centrándose fundamentalmente en promover la elaboración y difusión de materiales didácticos no sexistas, favorecer la igualdad de oportunidades en la promoción escolar, sensibilización de profesores y medio escolar en general que favorezcan un cambio de actitudes; c) en estrecha relación con el anterior, el campo de la formación profesional que ha tenido un gran empuje sobre todo con la creación de la red IRIS; d) en el sector laboral donde se estimulan un amplio abanico de acciones: apoyar las iniciativas locales de empleo, campañas de sensibilización, etc.; e) en el ámbito de las nuevas tecnologías, protección y seguridad social (Vighi, 1990; Instituto de la Mujer, 1991).

Los llamados estudios evaluativos que se han realizado sobre tales acciones ya sea a nivel europeo como a nivel de estado, son ejemplos ilustrativos de la confusión entre las realizaciones o productos —*outputs*— de una política y sus resultados o efectos, pretendidos o no —*outcomes*— y, en consecuencia, entre la gestión o auditoría y la evaluación. Así, en los documentos editados por los diferentes gobiernos o departamentos ministeriales bajo el título de evaluación de los planes de igualdad tratan de las acciones o actuaciones concretas que, en los diferentes sectores de intervención, se han llevado a cabo: número de proyectos de formación del profesorado subvencionados, publicaciones didácticas promocionadas o editadas, seminarios de formación y /o concienciación, revistas promocionadas o editadas, estudios de investigación sobre la problemática de la mujer subvencionados, etc. Así, debemos concebir más como informes de gestión que de evaluación los infor-

mes del Instituto de la mujer titulados: *Evaluación de la política comunitaria de igualdad de Oportunidades. Perspectivas futuras para 1992* y *Evaluación del plan de igualdad de oportunidades de las mujeres*; el informe presentado por el Principado de Asturias sobre *Evaluación del I Plan de igualdad de la Mujer*. En este sentido, es más realista el informe presentado por la Comissió Interdepartamental de Promoció de la Dona de Cataluña titulado como *Plan de actuación del Gobierno de la Generalitat de Cataluña para la Igualdad de Oportunidades para la Mujer*, puesto que en realidad se trataba de actuaciones concretas realizadas como resultado de la puesta en marcha del plan de oportunidades previsto para el período 1989-1992.

La investigación a gran escala dirigida por Subbarao y Raney (1993) responde a una política de concienciación sobre la educación femenina que promueve el Banco Mundial, además de ser un buen ejemplo de la utilidad de las macroinvestigaciones como inputs en la formulación de las políticas. Mediante un estudio de tipo correlacional con una muestra de 72 ciudades socioeconómicamente desarrolladas, evidencia que la reducción de la mortalidad infantil y fertilidad mantiene una relación sustancialmente mayor con el nivel de la educación femenina que la que tiene con los programas de planificación familiar y/o servicios de salud.

1.3. La utilidad de los resultados de las evaluaciones en las decisiones políticas

Un aspecto relevante en relación con la evaluación de las políticas es la utilidad de los resultados de la evaluación, constituyéndose en un tópico de investigación, como así lo ponen de manifiesto las sesiones dedicadas al tema en los últimos volúmenes de *Evaluation Studies Review Annual*. Como señala Rist (1990), en la obra específica que edita el autor: *Policy and Program Evaluation: perspectives on Design and Utilization*, existe un consenso de que un diseño riguroso es condición necesaria para que los resultados de los estudios evaluativos se utilicen, pero no es

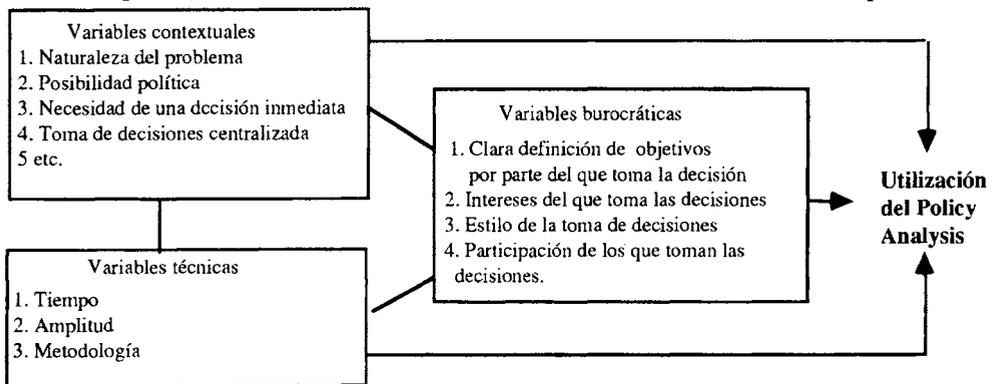


Figura 3

Un modelo conceptual de Lester sobre la utilización de resultados en los estudios políticos.

suficiente garantía. La cuestión bajo estudio es el cuándo y por qué los que toman decisiones políticas usan los resultados de las evaluaciones.

En la figura 3 se recoge la propuesta de Lester (1990) sobre un modelo conceptual generador de posibles hipótesis de investigación.

1.4. La investigación de indicadores en la mejora de los procesos de gestión pública

No hay duda que un campo importante en la investigación relacionada con las políticas públicas es la elaboración de indicadores que permitan, de un lado, controlar y estudiar la eficiencia de las políticas, y, de otro, de mayor interés para nosotros, valorar sus resultados e impacto. Se reconoce que una de las dificultades más importantes con las que se encuentra la evaluación de políticas es la falta de indicadores claros que permitan estimar la incidencia e impacto de tales políticas. Bartolomé (1992) refiriéndose al ámbito de la educación multicultural incluso llama la atención sobre la necesidad de elaborar indicadores que permitan determinar el nivel de interculturalidad de una sociedad determinada. En este sentido apunta el trabajo de Vega (1992) en sus esfuerzos por identificar indicadores que caractericen el nivel de «multiculturalidad» de una población.

Recientemente la OCDE ha publicado los resultados del trabajo que una comisión realiza sobre la elaboración de indicadores internacionales sobre educación: *The OECD International Education Indicators. A framework for Analysis*. En concreto, sobre los esfuerzos que se han hecho respecto al tema de los políticas y sus objetivos, los diferentes informes sobre indicadores (OCED, 1973; Unesco, 1976; Olkinuora, 1973; Netherlands Central Bureau of Statistics, 1988) contemplan un área específica de interés para el desarrollo de indicadores la de «igualdad de oportunidades» y la de «educación y calidad de vida»; áreas claramente relacionadas con políticas pretendidamente superadoras de situaciones educativas discriminatorias sea en razón del sexo, nivel socioeconómico o etnia. En esta línea debemos situar el reciente trabajo realizado por el CIDE sobre las desigualdades en el acceso a la educación en España.

Ya no como *indicadores de educación* —centrados en lo que ocurre en el sistema educativo— sino como *indicadores sociales de educación* —centrados en lo que ocurre en las condiciones sociales—, es loable el seminario sobre indicadores sociales promovido por el Instituto de La Mujer del Ministerio de Bienestar Social en septiembre de 1992. Consideramos de interés destacar la diferencia que establecen Saris y Batista (1992) entre *indicadores de proceso* (procesos subyacentes a una situación social) e *indicadores de estado* (situación del sistema). Los autores ilustran cómo determinar indicadores de proceso mediante un estudio sobre la discriminación sexual en el sistema educativo holandés. Identifican los consejos que los alumnos/as reciben de sus padres y profesores como indicadores de proceso que ayudan a explicar las diferencias que se observan en las ratios de matriculación en los niveles educativos, tomadas como indicadores de estado.

1.4. Los estudios de costes

Una preocupación cada vez más creciente en las actuaciones políticas se refiere a los aspectos económicos. Mitchell (1985) —citado por Grant (1992)— toma nota de ello en el caso de los programas multiculturales, si bien reconoce la pobreza de estudios sobre este aspecto. La reciente obra de Woodhall (1992) promovida por la UNESCO sobre análisis de costo-beneficios en la planificación de la educación, también es una muestra de ello.

Si bien se reconoce la necesidad de que las evaluaciones aporten algún tipo de análisis de costo, lo cierto es que las evidencias dicen que se utilizan muy poco. Smith y Smith (1985) en una revisión de la literatura encontraron muy pocos estudios evaluativos que utilizaron análisis de costos para hacer juicios comparativos sobre programas o intervenciones alternativas. En el campo educativo destacan los estudios realizados en los programas de educación compensatoria, si bien las revisiones realizadas por Wargo (1972) y la de Mullin y Summers (1983), ponen de manifiesto la práctica inexistencia de estudios serios de análisis de costes.

Smith y Smith (1985) dirigieron una encuesta nacional a los departamentos de educación de los EE.UU., a fin de describir los tipos de estudios de costo, las necesidades de modelos y los inconvenientes que se daban a nivel de estado. Los autores concluyen que: a) los métodos que se utilizaban eran los más simples; b) los principales impedimentos para hacer estudios de costos más serios y más útiles para la toma de decisiones eran la falta de preparación del personal responsable de hacerlo y la dificultad de relacionar costos con resultados; y c) en el campo educativo el análisis más prometedor son los estudios de costo-efectividad por la dificultad que entraña transformar en dinero los beneficios de programas educativos.

2. DIFERENCIAS ÉTNICAS Y SOCIOCULTURALES

Este apartado conecta con las síntesis presentadas recientemente en el Congreso de Educación Multicultural llevado a cabo en Ceuta (1991) y con el X Congreso Nacional de Pedagogía: «Educación Intercultural en la Perspectiva de la Europa Unida» (1992). De manera especial hemos conectado con la ponencia desarrollada por Bartolomé: «Diseños y metodologías de investigación desde la perspectiva de la educación intercultural».

Insistimos, sin embargo, en algunos de los problemas que plantea a la investigación la diversidad étnica y sociocultural, para presentar posteriormente los modelos más frecuentes utilizados en este tipo de trabajos tanto de enfoque cuantitativo como cualitativo.

2.1. Problemas que la diversidad étnica y sociocultural plantea a la investigación

La diversidad étnica y sociocultural plantea a la investigación educativa una

serie de dificultades y problemas. Una reflexión crítica sobre ellos nos lleva a ver que, aunque todos ellos puedan estar relacionados, son de diversa índole. Unos pueden influir en este tipo de investigaciones pero no son de tipo metodológico ni afectan a la validez de dichas investigaciones.

Así, por ejemplo, hay autores que han argumentado la dificultad de realizar este tipo de investigaciones en base a la falta de comprensión del significado de la educación multicultural y al etnocentrismo académico por parte de la comunidad científica. El primer hecho ha dado lugar a que las investigaciones que tratan de la diversidad étnica y sociocultural se hayan considerado estudios que sólo se refieren e interesan a los estudiantes de color, a grupos étnicos minoritarios, a poblaciones muy marginadas etc., y no a la mayoría de estudiantes. El hecho de que la comunidad científica pertenezca en su mayoría a la etnia y cultura dominante, tiende a favorecer una hegemonía cultural que monopoliza la construcción y legitimación de las perspectivas metodológicas al margen de las minorías étnicas. Este etnocentrismo académico se verá reflejado en los criterios empleados en la evaluación de trabajos de investigación o en lo que se ha llamado la «ghetorización» de las publicaciones periódicas. Por ejemplo, según Stanfield II (1993a, 12) muchas investigaciones sobre etnias en el ámbito anglosajón han sido galardonadas profesionalmente más por su conformidad con la ideología dominante que por su adecuación metodológica. También la ausencia de reflexión crítica sobre la metodología empleada puede ser una consecuencia del etnocentrismo cultural.

Otros problemas, los consideramos más relevantes pues producen sesgos que atentan a la validez y credibilidad de dichas investigaciones. Intentaremos hacer unas breves reflexiones sobre ellos y apuntar formas de evitarlos

1. El sesgo cultural del investigador

Autores como House (1990) y Hilliard (1989) indican que el sesgo cultural del investigador o del evaluador está implícito en las investigaciones, desde el planteamiento del problema hasta la interpretación de los resultados o las futuras recomendaciones de mejora.

La conceptualización de los problemas de investigación está precedida a menudo por sesgos ideológicos y culturales que determinan la producción de un determinado tipo de conocimiento de esa realidad. Autores como Prada (1992) señalan la importancia del sesgo del investigador cuando los resultados que ellos obtienen en sus estudios constituyen una fuente de información pública. Así, por ejemplo: en los cuestionarios sobre valoraciones de grupos étnicos no aparecen ciertos grupos étnicos mientras que siempre aparecen otros, reforzando así los estereotipos en la opinión pública hacia estos grupos; cuando analizan la discriminación, subrayan las actitudes discriminatorias más visibles y más alejadas de su propio contexto social, no percibiendo las más sutiles y enraizadas en su contexto.

Autores como Niles (1993) y Stanfield II (1993b) muestran la necesidad de reflexionar sobre ciertos aspectos relativos a la naturaleza de los problemas que se

investigan, la elaboración de las categorías y su relación con la identidad étnica y cultural, o la homogeneidad cultural.

a) La naturaleza de los problemas a investigar

Las cuestiones que se formulan en algunas investigaciones y las informaciones que se obtienen en ellas, tienen como finalidad documentar o poner de relieve problemas o patrones de aversión intergrupos (actitudes racistas, prejuicios, privación) más que relaciones de armonía y de cooperación. Aunque es lógico que se investiguen las situaciones de riesgo, este énfasis por resaltar los problemas entre etnias puede ser sintomático, según Stanfield II, (1993) de que muchos investigadores planteen sus investigaciones en el marco implícito de un modelo de asimilación.

b) La elaboración de categorías y la identidad étnica y cultural

Este tipo de investigaciones suelen implicar la categorización de personas, pero los criterios de diferenciación de estas categorías no suelen cuestionarse (Ogbu 1992). Suelen ser criterios objetivos, como el lugar de procedencia o rasgos físicos, ignorando otros criterios más subjetivos, como puede ser la identificación de los sujetos con el grupo étnico o cultural.

c) Homogeneidad cultural

Cada grupo étnico se trata como si existiera una homogeneidad cultural interna, atendiendo poco a la diversificación interna de cada categoría. En las investigaciones las variaciones intragrupo son tanto o más importantes que las intergrupo.

Los sesgos del investigador se reflejan, en el caso del observador, en la forma de percibir, seleccionar, registrar e interpretar una realidad. En este caso los investigadores actúan como filtro selectivo ante la realidad. Será preciso triangular observaciones y sus interpretaciones.

2. El problema de la comunicación

La comunicación verbal y no verbal es un serio problema en contextos de diversidad étnica y sociocultural. La distinta significación de los conceptos o de los gestos en las distintas culturas, afectará: a la comprensión de preguntas por parte de los sujetos, a la interpretación de sus respuestas por parte del investigador (por ejemplo en instrumentos: cuestionarios, entrevistas) y también a las relaciones interpersonales (relaciones profesor-alumno, interpretación de conductas por parte del profesor). En la elaboración de instrumentos es difícil proponer las mismas cuestiones y conseguir respuestas comparables.

3. Utilización de instrumentos sesgados culturalmente

En los estudios que investigan la ejecución diferencial en el rendimiento según la etnia de los sujetos, es frecuente el uso de tests estandarizados de rendimiento. Muchos autores nos han advertido del peligro que supone su utilización debido al sesgo cultural que conllevan. Así, por ejemplo, McDowell (1992) pone en entredicho la utilización de este tipo de tests cuando se evalúan programas educativos, debido a la ejecución diferencial. Además, nos dice, muchas de esas diferencias son atribuibles a factores que no son necesariamente mejorables por programas educativos, como pueden ser: la existencia real de desigualdad de oportunidades, las limitaciones técnicas de estos tests y la orientación cultural del postest (hacia tópicos, lenguaje y formas de conocimiento de unos grupo sobre otros). Nos advierte de la necesidad de crear tests culturalmente apropiados, y del peligro que hay si se utilizan en la evaluación de programas. «Los evaluadores tienen la necesidad y responsabilidad de comprender el papel que juega la cultura si desean practicar su oficio con responsabilidad y éticamente» (p. 52).

4. Trabajar con diseños y metodologías variadas

Si la diversidad étnica es una realidad compleja, también lo serán las investigaciones sobre dicha realidad. Stockdill (1992) señalan que en las evaluaciones de programas multiculturales, los métodos y diseños cuantitativos, así como determinadas técnicas de análisis estadístico son importantes, pero que deben complementarse con informes basados en el contexto del programa y que expliquen vivencias y experiencias de los participantes. Será necesario trabajar con diseños mixtos, combinando estudios extensos de campo e investigaciones etnográficas (Patton, 1987), y triangular las informaciones obtenidas a través de distintas vías.

5. Participación en las investigaciones de personas pertenecientes a las etnias a tratar

Algunos de los problemas mencionados surgen cuando los investigadores conocen muy poco acerca de los grupos étnicos sobre los que van a investigar, y por ello autores como Lincoln (1991), Madison (1992) y Stockdill (1992) señalan la necesidad de que personas procedentes de distintas culturas intervengan y participen tanto en la planificación como en la evaluación de programas educativos, cuando se trabaja con poblaciones de diversidad cultural. Madison (1992) afirma que ya que la evaluación es subjetiva al menos que sea más democrática e inclusiva, y que los propios interesados intervengan. Esta participación nos parece importante, si no queremos correr el riesgo de sesgar e invalidar las investigaciones.

En la investigación que estamos realizando en equipo en la Universidad de Barcelona desde 1992, bajo la dirección de M. Bartolomé, y titulada: «Diagnóstico de las diferencias étnicas y de los procesos desarrollados en la Educación Primaria»

se intentan paliar algunos de estos problemas: combinando estudios extensos de campo e investigación etnográfica, atendiendo a la realización de procesos educativos, introduciendo personas de las diversas etnias en la investigación, para que analicen críticamente los instrumentos contruidos, entrevisten a los padres de los niños pertenecientes a las minorías tratadas, y faciliten las interpretaciones de sus respuestas.

Pasamos a continuación a presentar algunos estudios y trabajos en intervención educativa, que traten sobre diferencias étnicas y socioculturales, agrupados según los modelos de investigación en que se basan y por áreas de contenido.

2.2. La investigación Ex-post-facto

Son los enfoques de investigación más abundantes. En el análisis de este tipo de investigaciones distinguimos dos amplios núcleos temáticos: uno de ellos orientado a comprobar los efectos de la educación en la diversidad cultural y el otro enfocado al campo de la orientación educativa.

Hay sin embargo muchos otros trabajos que aunque nos permiten conocer mejor a los diversos grupos étnicos, los hemos descartado por considerar que no eran fruto de una intervención educativa. Así, los trabajos destinados a analizar las diferencias étnicas en algún rasgo psicológico —por ej.: la percepción de agresión verbal (Phelps, 1991); o a comprobar alguna relación entre variables dentro de un mismo grupo étnico— por ej.: estudiar en niños negros el funcionamiento adaptativo y los comportamientos mal adaptados en relación a variables de tipo familiar (Dunn y Tucker, 1993), etc.

2.2.1. Estudiando los efectos de la educación en la diversidad cultural

Dentro de este núcleo temático, los estudios más abundantes se centran en analizar los efectos educativos en el rendimiento y logro académico, o bien los efectos educativos en las actitudes raciales, ya sean actitudes raciales entre alumnos, percepción de racismo por parte de los alumnos o lo que piensan los profesores de sus alumnos.

A. Los efectos de la educación en el rendimiento y logro académico

Aquí podemos distinguir entre investigaciones «input-ouput» (en la mayoría de los casos trabajando con poblaciones muy extensas), modelos estructurales, o estudios comparativos entre países. Algunos de ellos son estudios longitudinales.

Entre los estudios tipo «input-output» que prefieren trabajar con poblaciones extensas tenemos el de Kysel (1988) quién observa las variaciones que se producen en distintas cohortes y en distintos grupos étnicos sirviéndose de indicadores tan simples como el presentarse a examen o la promoción de nivel y observando las tendencias que se producen.

Otro estudio extenso ya más complejo lo tenemos en Rumberger y Willms (1992) quienes examinan la extensión de la segregación racial y étnica en las escuelas superiores de California y evalúan su contribución a las diferencias en el rendimiento entre los grupos étnicos. Como variables independientes trabajan con datos a nivel de estudiantes (la etnia, sexo, ejecución en inglés, nivel educativo de los padres) y a nivel de escuelas (variables contextuales y descripción del personal docente). Como variables dependientes, toman el rendimiento en lectura y matemáticas en los alumnos de los grados 3, 6, 8, y 12. Para medir la segregación elaboran dos índices de medida: un índice de disimilaridad y un índice de aislamiento, analizándose la segregación intra e inter escuela. Para estimar el impacto utilizan varios procedimientos: contrastes de medias, un modelo lineal jerárquico, técnicas de regresión multinivel y gráficos comparativos por grupos étnicos. De los resultados obtenidos en el estudio los autores destacan, entre otras cosas, la diferencia sustancial en el rendimiento entre los grupos raciales y étnicos, entre distritos y entre escuelas de los distritos; que los alumnos de las minorías pueden ir menos a escuelas con contextos escolares más favorables, y dar algunas orientaciones de reforma de política educativa tanto en la alteración de la composición social de las escuelas como en la redistribución de recursos. Como limitaciones del estudio se señalan la dificultad de separar los efectos del contexto de sus fuentes y el posible sesgo de sobrevalorar los efectos del contexto, pues al no haberse controlado el rendimiento de los estudiantes al entrar en la secundaria en las escuelas podían presentarse una acumulación de efectos tanto de la primaria como secundaria.

Otro estudio tipo *input-output* en el que se pretende comprobar el efecto del origen étnico y sus interacciones con el estatus socio-económico (SES) sobre el rendimiento, es el presentado por Tamir (1992). Establece 4 niveles de SES, creados a partir de un índice que incluye 11 variables, y como variables dependientes incluye —además de las puntuaciones en diversos tests— sus actitudes hacia las ciencias, intereses, aspiraciones educativas y ocupacionales, y autoconcepto. Para el análisis de datos utilizan técnicas de análisis multivariado.

Entre los *modelos estructurales* cabe resaltar la investigación realizada por Clifton (1991) quienes estudian el logro académico de grupos étnicos en estudiantes australianos, bajo un modelo teórico estructural psicológico social. Su finalidad es comprobar los efectos de la etnicidad tanto sobre el rendimiento académico como sobre su finalización de los 12 años de escolaridad secundaria. Los datos proceden de un estudio longitudinal de 7 años obtenidos entre 1975 y 1980, y en el análisis de datos utilizan métodos de análisis multivariados. Como puede verse en la Figura 1, el modelo asume que tanto las variables de estatus individual, como las variables académicas elegidas y las variables del contexto social, son todas variables exógenas. El rendimiento académico, la autoeficacia percibida y la influencia de otras personas significativas son variables intervinientes, y la finalización de los 12 años de escolaridad secundaria es la variable dependiente final.

Como síntesis de este trabajo podemos decir que los efectos de la etnicidad son relativamente pequeños, pero hay diferencias sustantivas entre los grupos étnicos.

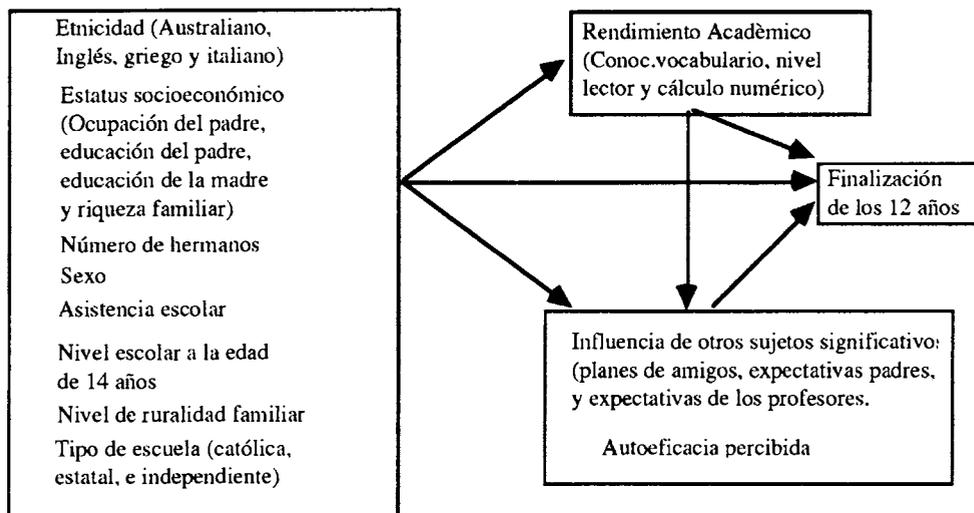


Figura 4
Modelo estructural de Clifton y otros (1991).

Los estudiantes griegos e italianos tienen un rendimiento académico inferior a los australianos e ingleses, el soporte social y psicológico que reciben para la educación es superior y hay mayor número de estudiantes griegos e italianos que completan la escuela secundaria. Estos resultados muestran que la ascendente movilidad de los estudiantes de habla no inglesa no está bloqueada en el sistema educativo. Más bien parece que el rendimiento de estos estudiantes está limitado a causa de su falta de ejecución en inglés, mientras que su éxito (al ser más el número de ellos que completan la secundaria), está facilitado por sus valores culturales (ejemplarizados en la parte inferior del gráfico).

Un ejemplo de trabajo extensivo en donde se hace un *estudio comparativo* entre los sistemas educativos de dos países es el realizado por Westbury (1992; 1993) y Baker (1993). Estos autores ejemplarizan diversos procedimientos cuantitativos para comparar rendimientos en matemáticas de estudiantes del Japón y EE.UU., comparando indirectamente la efectividad de los respectivos sistemas educativos a la vez que tratan el papel que juega el currículo. Los resultados muestran que aunque las clases de las escuelas americanas muestran más amplia variación en la ejecución anual, en Japón se enseña más matemáticas en la escuela elemental y secundaria, y se consigue mayor rendimiento.

Como indica Baker (1993), en la mayor parte de estudios comparativos sobre el rendimiento entre diferentes etnias ha habido un exceso de simplificación, sin analizar las causas que actualmente influyen en el rendimiento académico, ni las diferencias curriculares.

Dada la dificultad real que existe de realizar estudios comparativos cuando los

sujetos no siguen el mismo programa, y el escaso número de trabajos en que se compara la ejecución y los progresos de los diversos grupos étnicos, pero con programas educativos homogéneos, destacamos el de Power (1993). El estudio se centra en analizar las semejanzas y diferencias de los grupos pertenecientes a cuatro etnias en sus puntuaciones medias de grado (GPA) y sus puntuaciones en tests de aptitudes escolares (SAT) de todos los graduados de una universidad de Hawai desde 1975 hasta 1986. Se incluyen tres factores que pueden afectar a sus ejecuciones: etnicidad, sexo, y cohorte. Los efectos de estas variables se analizaron mediante el análisis de la varianza (ANOVA).

La conclusión de este estudio es que cuando los datos se toman de poblaciones escolares heterogéneas en un curso, emergen los modelos familiares de diferencia étnica. Cuando se examinan longitudinalmente por cohorte se evidencia poca variabilidad. En el grado 12 las ejecuciones en SAT entre caucásicos y asiáticos son similares (sobre estos dos grupos se han realizado muchos estudios comparativos). Es un estudio interesante porque nos advierte del peligro existente del exceso de generalización de las categorías étnicas en evaluación educativa. La evaluación educativa en una sociedad multiétnica debe ser sensible tanto a la variabilidad entre grupos como dentro de cada grupo; basarse en la comprensión y apreciación de las diferencias culturales y étnicas y debe evitar generalizaciones basadas en convenientes pero artificiales categorías de estudiantes.

Para la realización de estudios comparativos del rendimiento, puede ser interesante la aportación de Young (1991) desde el campo de la medición. Demuestra que las puntuaciones medias de grado acumulativas, basada en la teoría de la respuesta del ítem (IRTGPA) es un criterio mejorado en la predicción de las puntuaciones de preadmisión para los estudiantes pertenecientes a los diversos grupos étnicos (validez predictiva diferencial).

Si realizamos un análisis crítico de estos estudios que tratan sobre los efectos educativos en el rendimiento, vemos que en general:

— No investigan los procesos educativos diferenciados, sino la entrada y el producto diferenciado (diferencias en las variables de entrada y salida). Aún cuando se trabaje con estudios ex-post-facto, podría estudiarse el proceso.

— Utilizan tests de rendimiento que pueden dar lugar a errores por el sesgo cultural que conllevan.

— Son estudios, que como señala Ladson-Billings (1992), no contemplan otros resultados no cognitivos, como la mejora de relaciones casa-escuela o el aumento de la autoimagen de los estudiantes, es decir, no tienen en cuenta el impacto social y emocional del programa educativo sobre los estudiantes.

B. Los efectos de la educación en las actitudes raciales

La educación multicultural puede afectar profundamente a las investigaciones que tratan de estudiar precisamente los efectos de la educación en las actitudes raciales y que se realizan con poblaciones escolares. Las interacciones entre perso-

nas en contextos de diversidad étnica tratan principalmente las actitudes raciales entre alumnos, la percepción de racismo por parte de los alumnos en instituciones educativas o lo que piensan los profesores de sus alumnos.

Las actitudes raciales entre estudiantes se realizan con poblaciones escolares de diversas edades. Cuando se trabaja con niños pequeños es frecuente utilizar fotografías étnicas de niños desconocidos, que son validadas por jueces de las correspondientes etnias, y cuyo número varía según las variables que se desean controlar. Así, Ramsey (1991) estudia las preferencias de elección en niños de 3, 5, y 7 años a partir de 16 fotografías raciales igualadas en género, color de piel y diversas características fisonómicas. A los niños se les realizan distintas preguntas orientadas a la medición cognitiva y afectiva. En el análisis de datos utilizan el análisis de la varianza (ANOVAS), el contraste de medias «t» y la correlación de Pearson. Se comprobó que en la categorización de otros la raza es más frecuente que el sexo, mientras que en la categorización de sí mismos los dos rasgos sobresalían por igual; el sexo predominaba en sus elecciones positivas y la raza en sus elecciones negativas; los niños de mayor edad mencionaban con más frecuencia la raza como razón para rechazar a otros de diferente raza. La consistencia entre las respuestas cognitivas y afectivas aumentaban con la edad.

Estos resultados son congruentes con los estudios que versan sobre la formación de actitudes, y muestran la importancia de que los niños reciban una formación encaminada a evitar prejuicios desde pequeños. Por otra parte, el efecto sobresaliente de la raza ha sido bien remarcado por Davis (1992). El autor nos dice que no debe utilizarse este término porque acostumbra a ser mal entendido e interpretado, nos advierte del problema de utilizarla como una variable explicativa sobre todo de los efectos de un programa, y sugiere una serie de estrategias encaminadas a asegurar la adecuación de las interpretaciones.

Otra investigación parecida a la anterior, en la que se estudia el sesgo diferencial étnico en las elecciones emitidas por niños bengalíes y blancos, es el de Thirkell y Worrall (1989). Para estudiar los estereotipos étnicos hacen que las elecciones recaigan sobre grupos de niños más que sobre sujetos aislados.

Las actitudes étnicas/raciales se estudian también con poblaciones escolares mayores y en distintos ámbitos estudiantiles. Así, Morris (1991) mide en el campus universitario las actitudes raciales de los estudiantes que viven en viviendas griegas, y utiliza la chi cuadrado para comprobar en cada ítem de la escala las diferencias entre grupos, encontrándose diferencias significativas en todos menos uno. Las diferencias según el género se presentaban sobre todo en los ítems sociales. En el estudio de Sergent (1992) se miden las actitudes hacia los estudiantes árabes universitarios, y al igual que Engstrom y Sedlacek (1991) emplean la escala de actitudes situacionales (SAS) de Sedlacek y Brooks, compuesta por 10 situaciones sociales con escalas bipolares de diferencial semántico. Ambos analizaron los resultados usando el análisis de la varianza multivariado (MANOVA).

En cuanto a las actitudes de los españoles hacia los inmigrantes, el Centro de investigaciones sobre la realidad social (CIRES) recoge investigaciones realizadas

en 1991 y 1992, que permiten observar la evolución de las actitudes. Aunque los resultados no son fruto directo de una intervención educativa, pues se refieren a la población española, sí que nos orientan para futuras intervenciones. Así, por ejemplo resulta que los grupos étnicos peor valorados son los gitanos, seguidos de los árabes y judíos tanto en 1991 como en 1992 siendo, sin embargo, muy escasa la proporción de españoles que han mantenido algún tipo de relación, a nivel conversacional, con alguien de dichos grupos. Comprueban también que se ha producido un cambio de actitud, en cuanto a la constatación de los efectos negativos de la inmigración.

En relación a los resultados mencionados, cabe decir que autores como Ponterotto (1991), que han estudiado en profundidad los prejuicios y formas de combatirlos, muestran que el contacto personal es necesario para el establecimiento de relaciones positivas, pero si éste no se establece en unas determinadas condiciones puede reforzar los estereotipos existentes. Por otra parte, las actitudes racistas de los españoles y el tratamiento desigual que en nuestro contexto educativo tienen ciertas minorías y en especial la gitana, se destacan en otros estudios como los realizados por Calvo Buezas (1987; 1990). El hecho de que cuando en los textos escolares se habla de discriminaciones raciales y étnicas se sitúen siempre fuera de España puede crearnos la creencia ilusoria de que todos somos muy tolerantes y que nosotros, a diferencia de otros pueblos, estamos libres de prejuicios étnicos.

Para que se promuevan valores tales como el respeto a la dignidad humana y a los derechos humanos universales o el respeto a la cultura y la diversidad cultural, desde la Educación Multicultural es necesario contrarrestar las fuerzas existentes contrarias como el etnocentrismo, racismo, fascismo, la violencia o el vandalismo y quizás sea necesario para realizar una intervención educativa mejor elaborar instrumentos para medir la «tolerancia étnica» como los elaborados por De Lucas (1992) o por Sutter y McCaul (1993).

El reciente y progresivo aumento de niños inmigrantes en nuestras escuelas (no es hasta la década de los 80 cuando se empieza a hablar de España como país de inmigración) ha hecho despertar el interés en nuestra sociedad por la educación multicultural. Como indicadores del auge que está tomando en nuestro contexto las investigaciones educativas centradas en la diversidad étnica tenemos los numerosos estudios realizados por miembros del colectivo IOÉ o las nueve investigaciones financiadas por el C.I.D.E. en 1992/93 que están en curso de realización. En una de ellas intervenimos todos los participantes en esta ponencia bajo la dirección de Bartolomé (1992) y en ella se mide también las actitudes de los profesores hacia la educación multicultural. En dichas investigaciones intervienen grupos de personas distribuidos en una gran parte del territorio español.

Como notas características de las investigaciones educativas que sobre la diversidad étnica se desarrollan en nuestro contexto diremos que:

- Se realizan en equipo, y los equipos son multidisciplinares.
- El interés de dichos equipos por aunar esfuerzos y coordinar sus investigaciones, fomentando reuniones entre miembros de los distintos equipos.

— Algunas investigaciones se realizan de forma conjunta por distintas universidades.

Un ejemplo de este tipo de estudios, es el proyecto de investigación coordinado entre las universidades de Granada, Gerona y Murcia junto con la marroquí Mohamed V de Rabat, titulado «Conexiones entre la escolarización de hijos de inmigrantes marroquíes en centros escolares españoles (Málaga, Granada, Almería, Murcia, Barcelona y Gerona) y la escolarización en los centros escolares marroquíes de procedencia. Un estudio sobre las estrategias de adaptación e integración de niños inmigrantes y para el diseño posterior de un currículum multicultural». Este proyecto se presentó en 1992 y está financiado por varios organismos del Ministerio de Educación y Ciencia.

La percepción de racismo por parte de los sujetos de las minorías es importante, sobre todo si ésta tiene lugar en instituciones educativas, dedicadas a formar futuros educadores.

Un estudio interesante por sus repercusiones en la formación del profesorado es el de Siraj-Blatchford (1991) quien estudia las percepciones de racismo por parte de los estudiantes de centros de formación inicial del profesorado del Reino Unido. El estudio se refiere principalmente a 34 instituciones inglesas. El análisis de los resultados, aunque evidencian alguna debilidad en el diseño del cuestionario, —pues hay respuestas que denotan que alguna pregunta no ha sido bien comprendida— muestran el alarmante alto nivel de percepción de racismo. Así, tan sólo 5 de los 70 estudiantes no manifiestan evidencias de experiencias de racismo. El nivel de racismo en las escuelas, en algunos profesores y entre los compañeros, sugiere que se deberían desarrollar procesos de selección tanto del profesorado como de los estudiantes y que todas las instituciones que se dedican a la Formación del Profesorado deberían preparar a los estudiantes a enseñar en una sociedad multicultural, y establecer procedimientos para tratar con el racismo a todos los niveles.

Si el estudio que acabamos de mencionar muestra evidencias de racismo por parte de los profesores, percibidas por los estudiantes de grupos minoritarios, la investigación realizada recientemente en EE.UU. por Hlebowitst y Téllez (1993) con 235 profesores que están a punto de graduarse parece indicar lo contrario. El estudio pretende conocer lo que piensan estos estudiantes en relación a la raza, clase social y género, antes de interactuar como profesores en estos contextos. Ellos debían valorar el nivel de respeto que les producía cada una de las 8 descripciones que se les presentaba de estudiantes diferentes de educación primaria en las que se controlaban las variables mencionadas. Los datos de se analizaron mediante el Análisis de la Varianza (ANOVA) y el contraste de medias («t»). Resultaron, de forma significativa, que los niveles más altos de respeto eran hacia los estudiantes negros, con más bajo estatus social-económico y hacia las mujeres, aunque hacia este grupo resultó menos destacado.

Aunque los diferentes resultados de estos estudios pueden deberse a muchas causas, una de ellas puede ser que los sujetos de este último estudio no habían

tenido experiencias directas con niños, y es posibles que sus prejuicios emerjan y se hagan conscientes cuando estén en contacto con la realidad.

2.2.2. *La diversidad cultural y la orientación educativa*

El auge que está cobrando en EE.UU. el enfoque de la orientación educativa desde la perspectiva multicultural ha dado lugar a investigaciones en las que la valoración tanto del consejero como del propio proceso de orientación se asocia a variables como la etnicidad, nivel de aculturación o la identidad racial. Parece ponerse en evidencia que estas variables, y sobre todo el grado de sensibilidad cultural del consejero afectan a las relaciones en el consejo de orientación.

En la investigación de Gim (1991) se examinan los efectos que la etnicidad y la sensibilidad cultural del consejero y el nivel de aculturación y género de los sujetos, tienen sobre las percepciones de credibilidad y competencia cultural del consejero. Se utilizó diseño factorial $2 \times 2 \times 2 \times 2$ y los datos se analizaron mediante el Análisis de la Varianza Multivariado (MANOVA) con cada una de las dos variables dependientes. Los resultados del estudio sugieren que todas ellas influyen y juegan un papel importante en el grado como los estudiantes asiáticos perciben la competencia cultural y credibilidad del consejero.

Las preferencias de los consejeros también se han trabajado en relación a las actitudes de identidad racial. Helms y Carter (1991) recogen dos estudios en los que se examinan las relaciones entre las actitudes de identidad racial blanca o negra y la similaridad demográfica (clase social y género) sobre las preferencias del consejero. Los datos los analizan mediante el Análisis Jerárquico de Regresión Múltiple. Los resultados sugieren que la actitud de identidad racial opera de forma diferente en diferentes grupos raciales.

Bartolomé (1992) recoge las aportaciones de Sadowsky (1991) quienes presentan una interesante investigación multivariada sobre el efecto moderador de las variables socioculturales en los procesos de aculturación de dos grupos étnicos.

Para los orientadores que trabajan con poblaciones en transición transcultural les puede resultar muy útil conocer la serie de indicadores o descriptores del choque cultural, tanto positivos como negativos, así como las estrategias de intervención que aporta Zapf (1991).

Autores como Parham y Helms (1985a; 1985b) que han investigado sobre las actitudes de identidad racial en relación a otras variables psicológicas y su importancia en el consejo de orientación, sugieren que los americanos de origen africano en varios niveles de identidad racial pueden funcionar psicológicamente de forma distinta. Posteriormente, Carter (1991) valida en un estudio estas afirmaciones explorando cuándo las actitudes de identidad racial, según los 4 estadios desarrollados por Cross en 1978, eran predictivas del funcionamiento psicológico. Para el análisis de datos utilizan el Análisis de Regresión Múltiples.

También Helms (1984), con el ánimo de explicar los efectos de la raza en las relaciones de consejo, elabora modelos de desarrollo de la identidad racial, blanca y

negra. Otros investigadores como Tokar y Swanson (1991) pretenden validar alguno de estos modelos, en concreto el de identidad blanca, comprobando mediante procedimientos correlacionales uno de los supuestos básicos: que los estadios más avanzados estaban asociados a un mayor grado de autorealización.

Para finalizar este apartado señalamos otros estudios exploratorios de tipo descriptivo, que pueden ser útiles a los que trabajan en el campo de la orientación educativa, como son los de Parker y Lord (1993) que exploran las características de los modelos de rol en hombres jóvenes americanos de origen africano, utilizando la técnica de Análisis de los Componentes Principales; el estudio transcultural de Fredrickson (1992) en el que se estudian las semejanzas y diferencias entre los estudiantes de diferentes partes del mundo en la ordenación del estatus social de ocupaciones, o el de Lucas (1993) que recoge los problemas expresados por estudiantes de las minorías —tanto personales, sociales, académicos, y para la carrera— y en donde se pone de manifiesto la necesidad de realizar cursos de orientación específicamente diseñados para estas minorías.

En el área de la orientación educativa, encontramos alguna investigación en la que se utiliza un diseño pre-experimental (pretest-postest). Henry (1992) examina el impacto de un programa estructurado para la planificación de la carrera sobre las puntuaciones de madurez para la carrera en estudiantes americanos de origen africano. El impacto del programa se analizó contrastando las medias, antes y después, mediante la «t» de Student para muestras dependientes. Las diferencias entre chicos y chicas se realizó en el postest, con el mismo estadístico mencionado pero para muestras independientes. Los resultados mostraron, tanto en cada una de las escalas como en el total, diferencias significativas entre el pretest y postest. Dado que se cree que los Americanos de origen africano son más inmaduros para la carrera que los blancos y hay evidencias de que hay muy pocos en esta carrera, es necesario incorporar estrategias de intervención para este desarrollo a través de un proceso educativo. Es de esperar que el programa influya en un aumento de estos profesionales. La mayor limitación de este estudio es lógicamente la ausencia del grupo control.

2.3. La Investigación Experimental

El número de investigaciones experimentales es muy inferior al de los estudios ex-post-facto, quizás por la dificultad real que entraña el controlar las variables independientes. No hay muchas ocasiones en las que se pueda realizar la asignación aleatoria de los sujetos o de los tratamientos.

Aunque escasos encontramos algún trabajo interesante en el campo de la orientación, con investigaciones que someten a sujetos de diversa etnicidad a situaciones experimentales con la finalidad de comprobar los efectos de la aculturación en sus preferencias del consejero o comprobar el grado en que les afecta la sensibilidad cultural de éste.

Así, Hess y Street (1991) comprueban el efecto de la aculturación, género y

etnicidad del consejero, en la valoración de su efectividad. Los sujetos vieron 6 sesiones de vídeo en las que se simulaba una sesión de counseling habitual, en todas ellas tanto el consejero como el cliente eran varones y tan solo variaba la etnicidad. Se realizó un Diseño factorial $2 \times 2 \times 2$ para examinar las 4 variables dependientes que eran las puntuaciones dadas por los sujetos a las 4 dimensiones del Counselor Effectiveness Rating Scale: pericia, veracidad, atractividad y disponibilidad. Las variables independientes incluyen: etnicidad del consejero, nivel de aculturación del sujeto y género. Los resultados muestran que no hay diferencia significativa en los rangos asignados por los estudiantes Amer-Mejicano a los consejeros, según fueran de origen mejicano o inglés. Esta falta de preferencia por un consejero de su misma etnia también aparece en otros estudios, citados por los autores. Tampoco hubo diferencias significativas entre las preferencias de los hombres y las mujeres y sin embargo otros estudios indican que las mujeres tienden a valorar más la cultura inglesa.

A pesar de que en este estudio no se ha demostrado una valoración diferencial basada en el nivel de aculturación, sus autores señalan la importancia de que en los estudios de preferencia étnica intervengan como variables independientes variables intragrupo, como es en el presente caso la aculturación. La preferencia parece estar más marcada con el compromiso cultural.

Un estudio experimental interesante en el que se muestra con mayor claridad como la sensibilidad cultural puede ser un factor muy importante en las percepciones del consejero y en el proceso de Counseling, es el de Wade y Bernstein (1991) donde se examinan los efectos de un breve entrenamiento en sensibilidad cultural por parte de los consejeros y de su raza en las percepciones de los clientes (mujeres negras) de las características del consejero, la relación en el consejo y la satisfacción de los clientes con la orientación. El experimento se realiza con mujeres orientadoras (4 blancas y 4 negras), igualadas en cuanto a la edad, años de experiencia y nivel de dogmatismo. El grupo experimental recibió cuatro horas de instrucción, basado el modelo de asesoramiento transcultural de Pedersen (1985). La muestra de clientes estaba constituida por 80 mujeres negras de una amplia área metropolitana que respondían a unos anuncios del periódico o de la radio. 40 se asignaron al grupo experimental y 40 al grupo control. Tenían que valorar al consejero después de cada una de las tres primeras sesiones. A cada consejero se le asignaron 10 clientes, y de éstos se controló: ingresos familiares, edad, nº de miembros de la familia, años educativos, y condición familiar. Los datos se analizaron mediante el Análisis de la Varianza Multivariado (MANOVAs). Aunque el estudio es rico en resultados, quizás su mayor hallazgo sea la comprobación de que las percepciones de las clientes negras del consejero y del proceso de orientación, estaban más afectadas por la instrucción en sensibilidad cultural de los consejeros que por la raza del mismo. Además los clientes del grupo experimental volvían en mayor número a continuar las sesiones y mostraban un mayor grado de satisfacción con el proceso de orientación.

Este estudio apoya la hipótesis de D.W. Sue (1981) quien señala que cuando los

grupos son culturalmente distintos, la credibilidad del consejero se percibe en función del conocimiento que éste tenga de un grupo étnico específico.

2.4. Modelos cualitativos

En los modelos cualitativos el investigador se constituye como el instrumento básico de recogida de información, bien sea como etnógrafo, como observador participante, entrevistador o como analista de documentos. Consciente de que el núcleo principal de la información recogida está formada por las percepciones y vivencias subjetivas de las personas estudiadas, el investigador invierte mucho tiempo en lograr interacciones adecuadas (*rapport*) con las personas y en procurar que sus propias cargas valorativas no influyan en el proceso de investigación. A este respecto autores como Stanfield II (1993a, 8) ofrecen las siguientes consideraciones:

a) En una sociedad etnocéntrica no es sorprendente ver que, históricamente, las comunidades científicas de las etnias y culturas dominantes han elaborado las estrategias de recogida e interpretación de la información desde su propio marco cultural.

b) Hay que cuestionar el hecho de que investigadores de la cultura dominante puedan penetrar en la intersubjetividad de las personas pertenecientes a otras etnias. En caso de admitir esta posibilidad habría que determinar qué estrategias deberían seguirse para minimizar el inevitable sesgo que conlleva pertenecer a la etnia dominante.

c) Escribir parte de los informes de investigación o extractos de notas de campo en el lenguaje de culturas marginadas suele ser criticado en la comunidad científica de la etnia dominante. El lenguaje científico suele ser más conservador, menos afectivo y vivencial. En cambio, en el ámbito de las diferencias étnicas y socioculturales, las percepciones y vivencias expresadas en el mismo lenguaje de las etnias marginadas favorecen la comprensión más en profundidad del etnocentrismo (Stanfield II, 1993a), y permiten que los «excluidos de la historia» expresen su manera de concebir la realidad (Bartolomé, 1992). En este sentido, es cada vez es más frecuente encontrar informes etnográficos, especialmente estudios de casos como el de Harry (1992), en los que se recurre al lenguaje de minorías étnicas como pueden ser los destinatarios de programas de intervención familiar en el ámbito de la participación de los padres en la escuela. Esta posibilidad permite recoger en toda su viveza los conflictos entre el grupo cultural dominante y la etnia minoritaria, ofreciendo una visión más vívida de la conflictividad que generan muchos programas de intervención diferenciada.

En la literatura de investigación cualitativa que se ha revisado predominan los modelos de orientación *etnográfica*, y de una manera más incipiente, algunas derivaciones de la *investigación-acción*, aunque que no siempre logran cristalizar en un proceso real de optimización de las prácticas educativas interculturales.

2.4.1. Investigación etnográfica

Un objetivo fundamental de la observación participante es la descripción de grupos sociales y escenas culturales mediante la vivencia de las experiencias de las personas implicadas en un grupo o institución (Woods, 1987), con el fin de captar cómo definen su propia realidad y los constructos que organizan su mundo. De ahí que, en sentido amplio, autores como Jorgensen (1989) y McCall y Simmons (1969) conciben la observación participante, como una combinación de estrategias. Así, la observación directa de eventos relevantes ha de realizarse durante la interacción social en el escenario con los sujetos del estudio, unida a entrevistas formales e informales, registros sistemáticos, recogida de documentos y materiales, de forma flexible según la dirección que toma el estudio.

Los modelos de investigación con orientación etnográfica revisados se centran en ámbitos como el *prejuicio racial y étnico, cultura, comunidad y familia, clase social, identidad y actitudes políticas, integración racial y análisis de materiales y recursos*.

En el ámbito del prejuicio racial y étnico algunas investigaciones etnográficas han puesto de manifiesto que en los centros educativos puede existir racismo entre el profesorado (Carrington, 1983; Carrington y Wood, 1983 y Wright, 1986). Sin embargo, según (Foster, 1990) la evidencia de actitudes negativas del profesorado, en detrimento de grupos étnicos minoritarios, debe examinarse críticamente ante la posibilidad de ciertas deficiencias metodológicas, como pueden ser la descontextualización de las entrevistas y la falta de representatividad de los casos concretos.

En algunos estudios sobre *cultura, comunidad y familia* se evidencia la necesidad de entrevistas grupales (Lewis, 1992) con entrevistadores que puedan hablar correctamente la lengua del grupo minoritario como ocurre en el estudio de casos etnográfico con familias puertorriqueñas realizado por Harry (1992). A partir de su experiencia directa en investigaciones etnográficas Weis (1992) analiza la problemática que se plantea el investigador en un ámbito multicultural que, con frecuencia, resulta muy alejado del propio contexto cultural de procedencia.

La complejidad que encierran algunos ámbitos como el de la *identidad personal, las actitudes políticas y las diferencias culturales* exigen triangular información procedente no sólo de escalas y cuestionarios (Ghosh y Tarrow, 1993), sino también de entrevistas, individuales y grupales (Quicke, 1991), y de otras estrategias como pueden ser las historias de vida (Norquay, 1990), con el fin de captar mejor la dimensión afectiva actual a partir de la realidad pasada. Algunas investigaciones no sólo describen las características de estos constructos, sino también su *proceso de formación*. Así, autores como Weis (1988a, 1988b y 1989) analizan cómo se desarrolla la *identidad* en jóvenes de ambos sexos pertenecientes a la clase trabajadora y que asisten a un centro de secundaria enclavado en una zona industrializada de Estados Unidos. En otra investigación (Weis, 1985a y 1985b) aborda el desarrollo de la cultura negra en estudiantes que asisten a un colegio universitario urbano donde predominan alumnos/as con deprivación socioeconómica.

Según autores como Woods y Grugeon (1990) la investigación etnográfica es la más adecuada para describir los procesos de *integración y desintegración racial* entre escuelas, en las relaciones entre alumno/as y a nivel curricular.

Las investigaciones sobre *análisis de materiales y recursos* suelen centrarse en libros de texto y enfatizan el etnocentrismo de los libros de Historia. Desde una perspectiva didáctica Denia, Giró y Navarro (1988) analizan críticamente el etnocentrismo de los manuales escolares de Historia. A partir de imágenes e ilustraciones y de fragmentos de libros de las principales editoriales en nuestro contexto cultural, examinan críticamente aspectos como el nivel de etnocentrismo y de racismo, los estereotipos y sesgos etnocéntricos al considerar información y acontecimientos «importantes». Otras investigaciones como la de Maw (1991), citada por Bartolomé (1992), examinan los niveles de expresión directa, indirecta y de exclusión.

Algunas directrices metodológicas sugeridas por Bartolomé (1992) para optimizar los modelos etnográficos, están siendo aplicadas actualmente en la mencionada investigación (Bartolomé, 1992b) subvencionada por el CIDE y llevada a cabo por un equipo multidisciplinar. Este ambicioso proyecto nace con tres intenciones básicas a cubrir a largo plazo: ofrecer directrices para planificar una educación intercultural más allá de los modelos de integración o segregación; elaborar unas metas de educación intercultural y estrategias que contemplen las características y necesidades de esta población, incorporando avances y logros que ya comienzan a experimentarse desde la propia práctica reflexionada, y por último, crear una dinámica de formación de profesores para optimizar su propia práctica educativa intercultural.

La información necesaria para una adecuada planificación educativa intercultural exige un estudio comprensivo que permita clarificar qué *modelos educativos* se desarrollan en escuelas con pluralismo étnico y sociocultural, tratando de describir cómo se tratan en los centros las *necesidades educativas* de esa población infantil y qué *dificultades* se plantean los padres, el profesorado y el alumnado. Así, el primer objetivo consiste en realizar un diagnóstico de la población infantil inmigrante, étnicamente diferenciada, escolarizada en la Educación Primaria de la provincia de Barcelona. En segundo lugar, comprender en profundidad las relaciones educativas que se establecen en escuelas con diversidad étnica y sociocultural. Ambos objetivos se están cubriendo paralelamente, el primero a través de un estudio de campo, basado en una encuesta, y el segundo, mediante una investigación etnográfica.

La investigación etnográfica se está llevando a cabo en tres Colegios Públicos de Barcelona, seleccionados inicialmente según los criterios de accesibilidad y diversidad étnica y sociocultural, a partir de la información obtenida a través de contactos personales basados en amistades previas, con personal de la administración y con profesionales en ejercicio. Se optó por una estrategia de entrada abierta/directa por estimar que suele provocar, a la larga, menos problemas éticos y de autoconcepto en el observador, y por ser una estrategia más fácil, posibilitando un acceso más adecuado al fenómeno multicultural.

El proceso de observación participante consta de las fases descriptiva, focalizada

y selectiva (Spradley, 1988), adoptándose roles participantes diferenciados según lo exigen las circunstancias específicas en cada centro. Así, en uno de los centros la dirección y el profesorado mantiene una firme actitud de no reconocer la posible existencia de problemáticas en relación a las minorías étnicas y restringió inicialmente el proceso de observación lo que obligó a un largo proceso de negociación para poder acceder a las aulas. En consecuencia, la modalidad de acceso ha sido múltiple recurriéndose a la combinación de tres estrategias básicas:

— *Relaciones recíprocas*. Intercambio discreto de aportaciones, ayudas y beneficios que se intercambian. Así, se facilitaron pruebas de rendimiento a la Jefe de Estudios, un libro a la profesora de Compensatoria y realizamos sustituciones en las clases de profesores que tenían que asistir a Cursos de Reforma (se evitó así que los alumno/as se repartieran entre otros profesores/as).

— *Avalador reconocido*. La profesora de Compensatoria nos presenta a profesores/as.

— *Introducirse con un rol*. En general, dado que inicialmente no había una clara apertura a la observación y cooperación optamos por desempeñar un rol de personas que necesitábamos tener experiencia en el trato directo con las minorías étnicas estudiadas. Para esta finalidad proponíamos estar presentes en clases menos susceptibles de ser distorsionadas y que pudieran propiciar un mayor grado de participación, por ejemplo, dibujo o manualizaciones.

A lo largo de estas sesiones hemos recorrido el continuo que va desde una perspectiva completamente externa hasta una perspectiva completamente interna o participante. Incluso durante una misma sesión hemos tenido que adoptar diferentes grados de participación. Las cinco primeras sesiones en cada escenario se realizaron en común a cargo de dos observadores independientes, con el fin de ser analizadas por el conjunto del equipo, estimar la fiabilidad inter-observadores y clarificar las discrepancias. Posteriormente la información recogida durante la fase descriptiva se ha informatizado con los programas AQUAD y ETHNOGRAPH.

A partir de la experiencia vivida en la mencionada investigación etnográfica ofrecemos algunas consideraciones metodológicas referidas a la necesidad de utilizar múltiples perspectivas, conjugando adecuadamente la implicación y las posibilidades de observación. También es necesario clarificar conceptos clave, y optimizar la recogida de información con un adecuado control de la información.

Los procesos de *triangulación* (Taylor y Bogdan, 1986 y Marshall y Rossman, 1989) se han visto favorecidos con las estrategias de equipo, ya que facilitan la posibilidad de observar e interpretar la información desde diversas perspectivas, como son status social, sexo, edad, formación o experiencia y permiten establecer nuevas relaciones interpersonales, ampliando las posibilidades de observación. La posibilidad de ejercer *múltiples roles* también ha favorecido la observación desde distintas ópticas, aportando una visión más comprensiva y objetiva de la realidad étnica y multicultural.

La doble dimensión, *implicación* por un lado, y de *observación* por otro, ha de combinarse adecuadamente. La participación aporta percepciones y vivencias de las

personas implicadas desde una óptica interna. La dimensión de observación capta la realidad más a distancia o desde «fuera». Durante nuestra permanencia en los centros se plantea constantemente el reto de combinar de una manera óptima implicación y observación para ser capaces de reconstruir la realidad étnica y multicultural desde «dentro» con la intención de describirla también para los miembros del equipo que se sitúan en una perspectiva más «externa».

Durante las entrevistas informales surgen *conceptos clave* como pueden ser marginación, estereotipo o discriminación. Al definir dichos conceptos es conveniente realizar descripciones detalladas teniendo en cuenta el significado que las personas implicadas atribuyen a las ideas y acciones en situaciones particulares. Al formular conceptos basados en la perspectiva de los participantes cada observador participante ve facilitada la búsqueda de indicadores múltiples, incluyendo cómo se utiliza o se aplica el concepto en la realidad. De esta manera se pueden detectar con más facilidad posibles conexiones, asociaciones o relaciones entre conceptos. A este respecto, Jorgensen (1989) sugiere una posible estrategia como la siguiente:

- Elaborar inventario de palabras clave utilizadas por las personas implicadas.
- Analizar estas palabras pidiendo a dichas personas que describan en qué sentido se utiliza la palabra y observando cómo se utiliza.
- Generar nuevas observaciones o cuestiones orientadas a lo que se incluye en el dominio lingüístico, poniendo de relieve qué aspectos son similares o diferentes a partir de los símbolos o palabras utilizados.

La posibilidad de sesgos ideológicos y culturales en los informantes obliga a extremar el *control de la información*. Las entrevistas deben efectuarse según un proceso en espiral o de lanzadera que consiste en recoger información, analizarla y utilizarla como punto de partida para nuevas conversaciones e indagaciones. Cuando el entrevistador sospecha la posibilidad de sesgos puede recurrir a técnicas de *confrontación* y de *ruptura*. Por ejemplo, puede comprobar la veracidad de la información, confrontándola con otra aportada anteriormente por el mismo entrevistado, solicitando la clarificación de posibles discrepancias o contradicciones.

Al referirse a la *evaluación de programas* y a la *evaluación institucional*, Bartolomé (1992a) señala que en los países occidentales, muchas minorías étnicas constituyen grupos culturales marginados. En consecuencia, los programas de *educación compensatoria*, orientados a superar discriminaciones debidas al handicap sociocultural, han sido aplicados a estas minorías (Mira y Roig, 1991 y Codina, 1991). Según Bartolomé (1992a) los programas de *educación intercultural* (Aguado, 1991) y otros movimientos internacionales como son «La educación para la responsabilidad social» (Berman, 1990) o para la ciudadanía global o mundial (Sobol, 1990; Ramler, 1991 y Linch, 1992), constituyen un claro exponente del énfasis que se concede a la responsabilidad social, unida a la conciencia de pluralidad que caracteriza a la familia humana.

Bartolomé (1992a), retomando las aportaciones de Espín (1991) y de Linch (1992, p. 86) destaca algunas sugerencias para optimizar los modelos evaluativos:

—Necesidad de utilizar un modelo ecológico que permita la actuación desde diversos ámbitos y no se reduzca a la escuela.

—Reconocer que la influencia del profesorado es decisiva, por lo que urge vincular los procesos formativos del profesorado a los procesos de cambio educativo.

—Comprensión del problema dentro de un enfoque más amplio, de manera que la acción educativa se oriente a beneficiar a todos, entre ellos a los desfavorecidos.

—Necesidad de auto-evaluación profesional e institucional que contemple una supervisión democrática, un concepto emancipatorio de ciudadanía y los peligros de una endogamia moral y cultural.

Esta misma autora, recogiendo los trabajos de Knapp (1979) y Suttle (1986) comenta algunas propuestas suscitadas por la investigación evaluativa de orientación etnográfica, como son la necesidad de que el etnógrafo posea una sólida formación interdisciplinar y la conveniencia de no aislar la vida de la escuela de la comunidad que le sirve de referencia. En este sentido, tanto al recoger la información como al analizarla deben tenerse en cuenta *múltiples niveles contextuales*. Según Evertson y Green (1989) cada acontecimiento no sólo ha de observarse considerando el contexto inmediato, sino que debe insertarse en contextos más amplios que, a su vez, pueden influir en él. Entre todos ellos pueden existir influencias mutuas y redes de implicaciones que ayudarían a comprender e interpretar por qué ocurren determinados fenómenos. Además, tanto los *fenómenos* observados como el *escenario* donde se producen, deberían enmarcarse en su *contexto histórico* respectivo, viendo la evolución que han podido sufrir a lo largo del tiempo, lo que permite clarificar qué fenómenos son estables a lo largo del tiempo y cuáles son variables. Como es evidente, si el observador registra fenómenos relativamente estables obtendrá cierta representatividad. Sin embargo, la observación de otros fenómenos más variables, por ser más específicos de la situación o del momento temporal, ofrece menos garantías de representatividad.

2.4.2. Investigación-acción

Según Stanfield II (1993b) hay mucha necesidad de recurrir a modelos de investigación derivados de la investigación-acción que ayudarían a las personas implicadas a mejorar su cualidad de vida y a los profesionales de la educación a optimizar su intervención. Las posibilidades reales de estos modelos son analizadas ya por Bartolomé (1992a) con motivo del Congreso de Educación Multicultural, celebrado en Ceuta, y de la revisión de sus propias aportaciones (Bartolomé y Acosta, 1991; Bartolomé y Anguera, 1991 y Bartolomé, 1991 y 1992) y de otros autores (Anderson, 1989; Grugeon y Woods, 1990; Claux y Lemay, 1991 y Mateos, 1992) pueden extraerse valiosas orientaciones.

Algunas áreas que aparecen en la literatura de investigación como más susceptibles de modelos basados en la investigación acción son la *intervención familiar*, el *counseling*, *aculturación*, *rendimiento escolar*, *formación de profesores* y *programas de educación compensatoria*.

En el ámbito de la *intervención familiar*, la participación de los padres ha sido objeto de investigaciones como la de Harry (1992). Sin embargo, aunque se trataba de un modelo de investigación etnográfica de orientación crítica, con familias puertorriqueñas que tenían niños en educación especial, las barreras impuestas por la cultura dominante dificultan que el proceso llegue a cristalizar en investigación-acción pero inician un proceso de cambio, si bien ejercen alguna incidencia en la participación de los padres en la escuela.

Algunos autores como Ramisetty-Mikler (1993) describen diferencias socioculturales entre las culturas americana, india y asiática, analizando las implicaciones de estas diferencias y aportando sugerencias y directrices para enfoques y técnicas a utilizar en el counseling, que podrían sentar las bases para implantar la investigación-acción. Por otro lado, la presencia cada vez mayor de alumno/as biraciales procedentes de familias mixtas plantea un reto a la intervención a través del counseling para paliar problemas y necesidades derivadas de la ambigüedad étnica y la falta de identidad cultural. Según Herring (1992), esta área de intervención exige que los profesionales del counseling se impliquen en procesos de innovación y mejora, ofreciendo valiosas implicaciones para esta intervención especial que constituyen verdaderas hipótesis de acción y deberían fomentar una mayor atención a esta población especial.

Otras aportaciones como las de Pate y Bondi (1992) y Niles (1993) describen los peligros de los estereotipos y aportan directrices y recomendaciones para minimizar el sesgo étnico y cultural. A partir de las estrategias también podrían iniciarse procesos de investigación-acción.

Partiendo de tres modelos de *aculturación*, aculturación lineal (linear acculturation), presión acultural (acculturative stress) y presión de la marginación (marginality stress) Zimmerman y Sadowsky (1993) proponen prácticas culturales para tratar la alcohemia desde el counseling.

En el ámbito del *rendimiento escolar*, Robinson (1992) plantea numerosas iniciativas para transformar las prácticas educativas con minorías escolares de alto riesgo en base a la comprensión y apreciación de las diferencias étnicas y culturales. Brown (1992) estudia el proceso de *formación de profesores* desde la adquisición de nociones de educación multicultural hasta su aplicación práctica en el aula. Para ello, combina la utilización de escalas y cuestionarios con otras estrategias como son la observación en las aulas complementada con entrevistas.

En la línea de *formación de profesores* podemos situar la investigación cooperativa de Montero-Sieburth y Gray (1992), realizada con siete profesores de una escuela urbana de secundaria con diversidad étnica y cultural (blancos, latinos y chinos) ubicada en Boston. El proceso, que incorporaba metodologías interpretativas y de investigación-acción de orientación cooperativa, propició un clima de trabajo en el que unos profesores aprendían de otros mientras, reflexionaban críticamente sobre aspectos concretos que les preocupaban en su propio contexto educativo. La recogida de información se basó fundamentalmente en entrevistas abiertas, individuales y grupales a profesores, alumnos, estudiantes y personal de administración, con el fin de triangular las fuentes de información.

En el ámbito de los *programas de educación compensatoria* Bartolomé (1992a) recoge investigaciones con minorías étnicas como la de Mateos (1992), que elabora un detallado plan de acción para dar respuesta a los problemas detectados. La planificación de la acción se hace a partir de un diagnóstico del conocimiento que tiene una comunidad de maestros de las características de la cultura gitana y de los problemas que viven los niños de este colectivo.

En general, los modelos basados en la investigación-acción plantean dificultades cuando el proceso va más allá de la fase diagnóstica, ya que la espiral del cambio que anima el movimiento transformador requiere la implantación del plan de acción y la reflexión sobre sus efectos para lograr una progresiva optimización. Algunas directrices que podrían paliar los problemas más significativos pueden ser (Bartolomé, 1992a):

— El facilitador/a ha de poseer una comprensión sólida del problema a investigar y estará familiarizado con las áreas implicadas, con el fin de focalizar mejor los problemas y detectar posibles sesgos al percibir la realidad.

— Es aconsejable iniciarse o revisar estudios de etnografía crítica sobre el tema para emprender adecuadamente la investigación-acción.

3. DIFERENCIAS DE GÉNERO EN LA INTERVENCIÓN EDUCATIVA: UN TEMA VIGENTE EN LA INVESTIGACIÓN

3.1. Diferencias de género y modelos de investigación

El sexismo en investigación fue reconocido por vez primera como un problema en la década de los años 70. Aunque antes de estos años algunos libros y artículos habían señalado la existencia de él, no es hasta la segunda mitad de los 70 cuando las críticas aparecen con una cierta regularidad. En los años 80 diversas organizaciones y publicaciones comienzan a adoptar reglas sobre el uso no sexista del lenguaje y más recientemente sobre los contenidos no sexistas. Sin embargo, el sexismo en investigación está aún poco trabajado. Eichler (1991) analiza los tipos de problemas sexistas que pueden servir como instrumentos para facilitar el reconocimiento y la eliminación del sexismo en investigación. También elabora unas interesantes orientaciones para reconocer el sexismo en la investigación existente y evitarlo en investigaciones futuras. Estas orientaciones se concretan en una «checklist» muy útil para el análisis de la investigación desde una perspectiva no sexista. En ella se recoge el tipo de problema analizado, su descripción (para cada una de las fases del proceso de investigación), ejemplos concretos y posibles soluciones del mismo (Eichler, 1991).

Otra crítica del sexismo en investigación viene dada por la duda sobre la posibilidad de objetividad en las ciencias. Las investigadoras feministas mantienen que la objetividad es, en principio, imposible de lograr, ya que la ideología y el «background» del investigador influye al proceso de investigación en sus diferentes etapas (Messing, 1983; Keller, 1985; Eichler, 1991). Se afirma que cualquier inves-

tigación, incluida la investigación en educación sobre la igualdad entre los sexos, se apoya en un marco conceptual o en una perspectiva, elegida entre otras, sobre la condición humana. Quien investiga, sin embargo, no siempre es totalmente consciente de sus preconcepciones (Biklen y Shakeshaft, 1985). Así, muchas investigaciones se fundamentan en determinadas premisas que constituyen prejuicios sexistas no siempre reconocidos conscientemente, pero que forman parte del bagaje cultural del investigador.

También Patricia Campbell (1981) ha analizado cómo las investigaciones se ven afectadas por sesgos sexistas y racistas a lo largo de todo el proceso, desde la elección del problema a investigar hasta el análisis de los datos. Señala que hay que actuar con cautela, no aceptando como marco conceptual aquél que intenta interpretar la realidad de las mujeres partiendo de la experiencia de los hombres. Las investigaciones realizadas desde una óptica feminista mantienen que el problema a analizar configura el método adecuado, y no a la inversa; cualquiera que sea el método elegido deberá ir apoyándose en los nuevos datos que necesariamente se obtendrán al considerar la experiencia de las mujeres (Rubio y Mañeru, 1989). Si se quiere saber algo más sobre la educación de las mujeres y no se abandonan algunos de los paradigmas existentes, contruidos desde una perspectiva exclusivamente masculina, los modelos teóricos resultantes serán inexactos, imprecisos y en ocasiones, faltos de veracidad (Jiménez Fernández, 1990).

Una obra interesante para comprender la investigación desde una óptica feminista es la de Reinharz (1992). A partir del análisis de diferentes investigaciones en el ámbito social, hace una revisión de las diferentes metodologías de investigación, a la luz de la teoría y práctica feminista. Muchas de sus críticas son fácilmente trasladables al ámbito educativo. Dado que las críticas feministas demuestran que los métodos positivistas producen conocimiento desde una vía androcéntrica, las investigadoras feministas argumentan que la *entrevista no estructurada*, la *etnografía* y la *investigación-acción* pueden ser una alternativa importante para la investigación no sexista en el ámbito social y educativo. En esta misma línea, diferentes autores (Belenky, Clinchy, Goldberger y Tarule, 1986; Gilligan, 1979; Harding y Hintikka, 1983 y Scher y Good, 1990) llegan incluso a afirmar que dado que el método hipotético-deductivo es un modelo masculino (lógico, racional, medible) mientras que el modelo empírico-observacional es femenino (intuitivo, no linear), éste último debería tomar más relevancia en la investigación en este ámbito.

Como consecuencia de lo anterior, a partir de los años 80, las investigaciones empezaron a basarse cada vez más en las formas de investigación cualitativas: entrevistas abiertas, investigaciones etnográficas, observación participante, historias reales y estrategias de investigación participativa. La atención se centró en tratar de comprender la realidad de la vida de las mujeres como base para entender por qué y cómo éstas eligen las formas de educación para sí mismas y para sus hijas (Kelly, 1989).

Un artículo interesante para conocer la investigación educativa sobre desigualdades de género realizada en Holanda desde 1982 a 1991 es el de Ten Dam y Volman

(1991). Lo recogemos aquí ya que la mayoría de los tópicos investigados también lo han sido en otros países, entre ellos el nuestro³.

Estas autoras agrupan estas investigaciones en dos grandes tipos: uno, que denominan *investigación orientada políticamente*, muy unida a la educación primaria y secundaria; y el otro, *investigación-acción* que se desarrolla en el ámbito de la educación de adultos. La primera está financiada por el Ministerio de Educación y Ciencia de este país mientras que la segunda es independiente del gobierno central y se inspira en el movimiento de la mujer.

La *investigación orientada políticamente* utiliza una metodología *cuantitativa* (Alting y Pelgrum, 1990) y toda ella está centrada en el concepto de «desventaja» sobre todo, en las desventajas de las adolescentes en las escuelas secundarias. Algunas de estas investigaciones concluyen que las desigualdades de género en educación no son una cuestión de diferencias de logros académicos sino de diferencias en actitudes e intereses producidos por los procesos de socialización.

Otros estudios se centran en el análisis de los factores que influyen en las elecciones educativas de las adolescentes (división sexual del trabajo, características de los alumnos, interacción profesor/alumnos, actitudes del profesorado y funcionamiento del propio sistema educativo). Se concluye que la reproducción de las desigualdades puede ser explicada por la existencia continuada de determinados elementos de las prácticas educativas tradicionales en relación al sexo femenino, que están explícitamente orientados a la reproducción de la desigualdad. En este proceso, se considera al profesor como un importante reproductor de estas desigualdades.

También ha sido objeto de investigación el por qué de la escasa participación de mujeres en matemáticas, ciencias y materias o profesiones técnicas, y su motivación para ello. El análisis del currículum, de los materiales de enseñanza, el papel del profesorado y la didáctica básica de la educación en el campo de las ciencias, ven incrementados su niveles de prioridad en la investigación de los últimos años, al ser considerados como posibles explicaciones para las elecciones y logros académicos de las adolescentes.

La *investigación-acción* ha evolucionado de estudiar el aprendizaje por la experiencia como foco principal, a estudiar lo que se denomina el «aprendizaje femenino». Las formas específicas que las mujeres como grupo ponen en juego en actividades de aprendizaje. Esta aproximación al «aprendizaje femenino» se ocupa de la revalorización y desarrollo futuro de las cualidades de la mujer.

3.2. Género, currículum y clima de la clase

Numerosas voces argumentan que el sexismo en el contexto educativo ha disminuido. Sin embargo, el análisis de las investigaciones realizadas en los últimos años en la escuela, pone de manifiesto que esta premisa es prematura (Lafrance, 1991;

3 Otro artículo interesante para conocer la investigación educativa sobre las diferencias de género en los países escandinavos es el de Scott, A. (1992). The Question of Representation: reseach in gender and education in Scandinavia. *Gender and Education*, V. 4. Nº 3, pp. 201-211.

Askew y Ross, 1991, Caselles, 1991). La investigación muestra que las desigualdades de género en la clase continúan, se manifiestan y se mantienen en una variedad de aprendizajes no conscientes muchas veces, mensajes verbales y no verbales, iniciados en las interacciones entre el profesorado y el alumnado. Este tipo de investigación intenta analizar o bien el currículum explícito (estudios referidos al análisis de los materiales escolares, sobre todo los libros de texto) o bien el denominado currículum oculto (que obliga a analizar las interacciones que se producen en el aula).

Los modelos de investigación sobre diferencias de género en este ámbito se han centrado básicamente en: investigaciones orientadas a la comprensión de los procesos de intervención desde una perspectiva diferencial (*estudios observacionales, análisis de contenido* de materiales escolares o libros de texto, *estudios etnográficos y estudios de casos*); investigaciones orientadas al cambio educativo (*investigación-acción*); estudios *ex post facto* y *causales* y algunas, las menos, investigaciones *experimentales o cuasiexperimentales*.

Los *estudios observacionales* —los más numerosos— se han utilizado para el análisis de las interacciones que se producen en el aula, centrándose sobre todo en la búsqueda del posible tratamiento desigual dado al alumnado. Así, el *estudio observacional* de Cospér (1970) encontró que el profesorado inicia más interacciones verbales con los niños, discrimina negativamente a las niñas y es más restrictivo con ellas; además los niños inician más interacciones verbales con los docentes que las niñas.

Hillman y Davenport (1978) (citados por Brophy, 1985) observaron la interacción entre el alumnado y profesorado de 306 clases de preescolar y primaria. Los resultados obtenidos indican que se establece mayor interacción con los niños y se les riñe más y que ellos preguntan más al profesor/a. No se encontraron diferencias en relación al sexo del docente, ni por el hecho de que el docente y alumno pertenezcan al mismo sexo. Por su parte, Sadker y Sadker (1985) *observaron* a alumnos de diferentes clases de escuelas primarias y encontraron que en todos los grados y en todas las materias, los chicos dominaban las comunicaciones de la clase; participaban en más interacciones que las niñas. También los datos de Brophy (1985) sugieren que la diferencia de sexo de los alumnos supone más diferencias en el comportamiento de la clase y en la interacción con el profesor que las diferencias de sexo entre los profesores.

Perdue y Connor (1978) codificaron, a partir de una *observación* previa, no solo la frecuencia de los contactos directos del profesorado de preescolar con sus alumnos sino que también categorizaron el tipo de contacto que realizaban. Codificaron cuatro categorías de contactos: amistosos, útiles, de atención e incidentales. Los resultados mostraron que, particularmente en el caso de los profesores del sexo masculino, las niñas recibieron más contactos «útiles» que de otros tipos; los niños recibieron más contactos «amistosos». Los autores interpretan estos resultados como indicadores de una conducta benevolente y protectora de los profesores hacia las niñas que puede conducir a unas relaciones de dependencia.

El estudio observacional de Fried (1989) investiga los procesos de intercambio lingüístico que se producen en el jardín de infancia entre profesorado y alumnado, a partir del análisis de conversaciones de grupo sobre juegos lingüísticos o libros ilustrados. Sus análisis cuantitativos de las frecuencias de intervención no pusieron de manifiesto diferencias globales relacionadas con el sexo. Sólo un análisis cualitativo más minucioso permitió descubrir la existencia de estructuras lingüísticas diferentes para niños y niñas. Esta autora concluye que se ofrece a las niñas un margen de aprendizaje lingüístico más estrecho y menos estimulante, lo que le lleva a afirmar que los niños cuentan con más posibilidades de mejorar su capacidad lingüística. Otras investigaciones (Kelly, 1991) han *observado* las interacciones lingüísticas que se producen en clases de adultos, volviendo a poner de manifiesto la preponderancia del género masculino sobre el femenino.

Otras investigaciones sugieren que el comportamiento del profesorado varía según la materia de que se trate (Wilson y Handley, 1986; Worrall y Tsarna, 1987); la influencia de la asignatura impartida aparece cada vez más como un dato importante a tener en cuenta en este tipo de investigaciones, ya que probablemente la interacción establecida en el aula varía en función de las expectativas de los docentes respecto a la mayor adecuación de los contenidos para niños o para niñas (Davinson y Kanyuka, 1992; Bennett y otros, 1993), también puede variar la interacción por la actitud de alumnos y alumnas según lo que creen que se espera de ellos.

En nuestro contexto, se han realizado también algunas investigaciones para analizar el sexismo y el tratamiento diferencial que se aplica al alumnado en la educación primaria, utilizando la metodología observacional (Subirats, 1985; Subirats, 1987; Subirats y Brullet, 1988).

El trabajo de Subirats y Brullet (1988) sobre la trasmisión de los géneros en la escuela mixta, profundiza los trabajos anteriores de Subirats (1985, 1987). Se analizan las relaciones verbales que se establecen entre el alumnado y profesorado, las opiniones de los docentes sobre las posibles diferencias de comportamiento y actitud de los niños y niñas y los comportamientos verbales del alumnado, a partir de las *observaciones y entrevistas* realizadas en 11 escuelas primarias de Cataluña. Estas autoras concluyen que se ha producido una casi total desaparición de las diferencias institucionalizadas, excepto en lo que se refiere al deporte y espacios más directamente relacionados con el cuerpo, y que se producen ciertas diferencias en las demandas explícitas a cada grupo sexual, pero no sistemáticas ni legitimadas.

Ahora bien, el análisis realizado sobre comportamientos verbales, que en parte escapan al control consciente del profesorado, muestra la pervivencia de notables diferencias. Éstas se concretan en una mayor atención a los niños, mayor atención que ha sido medida a través del número de palabras e interpelaciones dirigidas a ellos y ellas. De modo general, la relación obtenida es de 100 a 74 a favor de los niños. También encuentran que esta ratio cambia en función del contexto conductual. El análisis de la variable *curso* indica que la discriminación lingüística tiende a aumentar cuando se pasa de preescolar a primer ciclo de primaria, y tiende a

disminuir a partir del 6º curso. Las profesoras acentúan más que los profesores la discriminación lingüística en relación a las niñas, prestan mayor atención a los niños que los profesores.

El análisis de las *entrevistas semiestructuradas* con el profesorado muestra la persistencia de los estereotipos ligados al sexo aunque se apunta algún cambio en estos estereotipos. Las conclusiones de este estudio indican una cierta tendencia del sistema educativo hacia un tratamiento más consistente de los niños y niñas, independientemente de su sexo. Sin embargo, esta tendencia está basada en la universalización del modelo masculino y no en una fusión de los modelos femenino y masculino.

Se necesitan investigaciones más detalladas y longitudinales que revelen por qué las interacciones en el aula crean y/o mantienen las diferencias de género (Lafrance, 1991). El artículo de Lewis (1990) plantea interesantes sugerencias para transformar una clase (las interacciones que se producen entre el profesorado y el alumnado) desde la perspectiva feminista.

Las investigaciones que *analizan los libros de texto* —las más tempranas en el ámbito que nos ocupa— utilizando la metodología del *análisis de contenido*, estudian los siguientes aspectos:

- La frecuencia con que aparece la mujer en el texto y en las ilustraciones.
- El vocabulario empleado cuando se refieren al hombre o la mujer.
- El retrato de la mujer en diferentes contextos (familiar, escolar, participación política, vida profesional, etc.) tanto en el texto como en las ilustraciones.

(Álvaro y Monge, 1984; Garreta, 1984; Schweitzer, 1985; Careaga, 1987; Heras, 1987; Wilson, 1991; Plateau, 1991; Polydorides, 1991; Jiménez Aleixandre, 1992; Marco, 1992).

Todos estos estudios confirman que los libros escolares presentan un rango restringido y estereotipado de las mujeres y niñas como sujetos pasivos en la sociedad. Se afirma que este punto de vista distorsionado de la realidad, ha tenido una fuerte influencia en la autoimagen de los chicos y las chicas, hipótesis que necesita ser explorada en profundidad (Wilson, 1991).

El estudio de Shilling (1991) analiza el *uso del espacio escolar* como un recurso para la producción de relaciones de género desiguales. Se ve el espacio no solo como un contexto en el que ocurren las interacciones sino como un fenómeno que produce y a su vez es producido por relaciones de poder mediatizadas por el género. Se argumenta que el estudio del espacio debe ocupar un papel importante en los trabajos sobre desigualdades de género. En opinión de la autora, los estudios etnográficos de las escuelas y las clases deben tomar en consideración como elemento importante el uso que se hace del espacio escolar; es del todo necesario para poder interpretar adecuadamente los resultados de estos estudios.

Es interesante el *estudio etnográfico*, de dos años de duración, realizado por Crump (1990) en una escuela secundaria para explorar las prácticas de resolución de problemas de los alumnos, utilizadas en las relaciones sociales entre el alumnado y el profesorado, y para valorar la coherencia de ellas; prácticas relativas a las

cuestiones de género, poder y currículum. Las técnicas utilizadas son: entrevistas, cuestionarios, observación participante, discusiones de grupo y el análisis de documentos institucionales.

En el contexto europeo, llama la atención el contraste entre los resultados de las investigaciones que analizan las actitudes del profesorado hacia las cuestiones de género y el compromiso expresado por éstos respecto al principio de igualdad de oportunidades o a las estrategias de enseñanza prácticas que utilizan en su trabajo diario. El profesorado a menudo da un tratamiento diferencial a sus alumnos y alumnas, como ponen de manifiesto los diferentes *estudios de casos* realizados en diferentes países europeos (Wilson, 1991). Estos estudios resaltan la mayor atención que el profesorado dedica a los alumnos del género masculino, sobre todo en la formas de alabanza y censura y en sus preferencias expresadas por la enseñanza de los chicos, especialmente en los niveles superiores de la escuela y en ciertas materias como ciencias, matemáticas o informática. De acuerdo con tales estudios, los chicos explotan esta ventaja en las clases según sus intereses. También estos estudios señalan una marcada separación entre los sexos en las clases coeducativas.

Con respecto a la *disciplina de la clase*, el trabajo de Robinson (1992) investiga su relación con las actitudes del profesorado hacia los modelos estereotipados de masculinidad y feminidad. Se llevan a cabo *estudios de casos* en seis escuelas diferentes donde asisten alumnos desde preescolar hasta los 12 años. Las técnicas utilizadas son: cuestionarios, entrevistas individuales y grupales y observaciones en las aulas. Se concluye que las actitudes estereotipadas del profesorado hacia la masculinidad y feminidad se reflejan en sus métodos de enseñanza y en sus prácticas en la clase. Como esta investigación sugiere, la disciplina de la clase es un área importante donde tales actitudes florecen y donde pueden reforzarse en el alumnado los dobles estándares y valores tradicionales. Esto produce serios efectos sobre la calidad de la escolarización del género femenino, su habilidad para desarrollar sus potenciales y sus futuras elecciones vocacionales.

El estudio de Alberdi (1987) sobre las actitudes de los profesores de escuelas secundarias basado en un *estudio cualitativo* con una pequeña muestra de profesores de Madrid, mostró que los profesores en general aceptan que tratan por igual a chicos y chicas, independientemente de su sexo. Se utiliza el *grupo de discusión* como técnica metodológica por considerar «que, para conocer las pautas que determinan los comportamientos, es necesario reconocer el compromiso de la conducta individual con el grupo social» (p. 28). El análisis de las intervenciones del profesorado en discusiones de grupo mostró poca conciencia de la persistencia del sexismo; creen que éste ha cesado una vez que se ha establecido la coeducación. Sin embargo, en las observaciones que hacen de sus alumnos reproducen de forma estereotipada lo que la sociedad viene atribuyendo a los géneros masculino y femenino como rasgos intrínsecos de cada uno. Se percibe, por tanto, una aceptación general de los estereotipos.

Por otra parte, siguen siendo todavía abundantes las investigaciones cercanas al ámbito de la psicología diferencial que tratan de probar diferencias de género en

habilidades intelectuales (Hedges y Friedman, 1993); en conductas agresivas y valores (Horowitz, 1992); en el desarrollo sociomoral (Pérez-Delgado y otros, 1990); en la elección de determinadas materias escolares (Archer y McDonald, 1991) o en la ejecución diferencial entre chicos y chicas en algunas de estas materias, sobre todo en el campo de las matemáticas, ciencias e informática (Held y otros, 1993; Low y Over, 1993; Hall y Cooper, 1991); los estereotipos y actitudes de los alumnos hacia un rendimiento académico elevado de las chicas (Faulkner, 1991) y los sesgos de género en las técnicas utilizadas para la evaluación del alumnado (Young, 1991; Stobart, Elwood y Qianlan, 1992). La mayoría de estos estudios utilizan la metodología *ex post facto* para establecer estas diferencias entre los sujetos, objeto de estudio.

Otras investigaciones, utilizando el *análisis causal*, estudian la relación entre la autoestima académica y el autoconcepto general y académico de alumnos de ambos sexos con sus expectativas de logro en matemáticas e idiomas (Skaalvik, 1990; Skaalvik y Rankin, 1990) o bien las relaciones entre las expectativas de autoeficacia y las atribuciones causales para el rendimiento entre alumnos y alumnas universitarios (Vasil, 1992).

Diferentes trabajos actuales analizan las percepciones del alumnado de primaria de los estereotipos ligados al rol sexual. Así, Gombos-Tozzo y Golub (1990) experimentan una nueva técnica para modificar estos estereotipos con alumnos de tercer grado de primaria (8-9 años). Consiste en presentar al alumno, durante un determinado tiempo, una descripción de diferentes profesiones, seguida de la realización por parte de los alumnos de una actividad asociada a cada profesión. Las mismas profesiones se presentan tanto por monitores del sexo al que tradicionalmente se le adjudica dicha profesión, como por monitores del sexo contrario. Las autoras parten de la hipótesis de que los niños, realizando actividades con modelos que contradicen los estereotipos ocupacionales tradicionales, pueden cambiar sus percepciones estereotipadas. Se utiliza un *diseño experimental pretest-posttest* para probar los efectos producidos por la técnica diseñada. Encuentran efectos significativos que indican que tanto los niños como las niñas del grupo experimental modifican sus percepciones iniciales de los estereotipos ocupacionales femeninos. Los resultados sugieren que los modelos de interacción alumno/profesor pueden ser importantes mediadores de las diferencias de género en las expectativas del alumnado (Gombos-Tozzo y Golub, 1990, 128).

Por otro lado, hay una serie de estudios que evalúan el impacto de la experimentación con *materiales instructivos* que presentan al hombre y a la mujer en roles no tradicionales. Schau y Scott (1984) llevan a cabo una revisión de 21 estudios donde se examinan los efectos de estos materiales sobre las actitudes del alumnado hacia los roles en función del género. Llegan a la conclusión de que las actitudes de los alumnos que trabajan con estos materiales son menos estereotipadas aunque los cambios no se pueden generalizar ni a todos los contenidos ni a todos los materiales; no encuentran modelos consistentes. La evidencia sólo permite especulaciones sobre la persistencia de los efectos, los cambios que se producen según diferentes

grupos de edad o los efectos de la cantidad de tiempo durante el cual los alumnos están trabajando con dichos materiales.

El análisis del impacto de las *intervenciones curriculares* que exponen a los niños a modelos de rol de género no tradicionales, no permite generalizar los resultados a todas las situaciones educativas ni a todos los roles. Cuando se encuentran efectos positivos son siempre a corto plazo ya que los efectos a largo plazo no se han evaluado (Leming, 1992)

Por último, *el aprendizaje cooperativo* que ha tenido efectos positivos, facilitando las interacciones intergrupos y las elecciones interétnicas, parece tener menos impacto en las relaciones de género y los estereotipos (Leming, 1992). Sadker, Sadker y Klein (1991) han realizado una revisión al respecto y afirman que los limitados resultados de la investigación sugieren que el aprendizaje cooperativo no lleva a un ambiente de aprendizaje más equitativo para los sexos o promueve interacción entre sexos. Afirman que cuando se trabaja en grupos cooperativos, los modelos culturales de agresividad masculina y dominancia se asocian con los modelos tradicionales de inhibición femenina en contextos sociales y pueden convertirse en modelos tradicionales de interacciones de género. Señalan, por consiguiente, que los estereotipos raciales y étnicos pueden ser más cambiables que los estereotipos de género cuando se utilizan experiencias de aprendizaje cooperativo.

3.3. Formación del profesorado desde un enfoque no sexista

La investigación ha estudiado las relaciones de clase, sexo y género en la profesión de la enseñanza, analizando el por qué se da una proletarización de la misma a la vez que aumenta la proporción de mujeres en ella (Casey y Apple, 1992). A pesar de ser fundamentalmente mujeres las que llevan el peso de la enseñanza primaria y buena parte de la secundaria (Wilson, 1991; Calzada, 1987), ¿por qué permanece la orientación masculina en la escuela? Este y otros muchos interrogantes han llevado a plantear la necesidad y urgencia de formar al profesorado desde un enfoque no sexista (Bartolomé, 1991).

Los modelos de investigación utilizados en este ámbito se insertan en el enfoque cualitativo. Se utiliza sobre todo la aproximación biográfica y la investigación-acción.

Así, el estudio de Sikes (1991) se focaliza en la importancia de ayudar a los alumnos de los primeros cursos de magisterio a tomar conciencia de las experiencias educativas de las muchachas y muchachos y cómo algunos estereotipos de género son el resultado de sus propias biografías. Esta concienciación es esencial para que cuando sean profesores, puedan reconocer y trabajar en la eliminación de la discriminación y los sesgos que existen en las escuelas. Los datos empíricos aportados, obtenidos mediante una *aproximación biográfica*, se usan para demostrar por qué el género y las políticas sexuales deben estar explícitamente dirigidas en la educación inicial del profesor y porqué estos profesores deben tener habilidades de análisis e instrumentos conceptuales que les ayuden a reflexionar e interro-

garse sobre sus propias experiencias en su casa y en la sociedad. Esta autora también considera la necesidad de examinar los cursos de educación inicial del profesorado para evitar el refuerzo de los estereotipos ligados al rol del sexo.

Por su parte, el Instituto de la Mujer ha realizado dos encuentros sobre la formación inicial del profesorado en educación no sexista; las Primeras Jornadas en noviembre de 1991 y las Segundas en febrero de 1992. En estas jornadas se presentaron una serie de programas desarrollados en diferentes Escuelas Universitarias del Profesorado de nuestro país (Valencia, Melilla, Palencia, Logroño, Valladolid, Burgos, Soria,...) con la finalidad de integrar la igualdad de oportunidades en el currículum de la formación inicial del profesorado. Estos programas, subvencionados por la Comunidad Económica Europea, se integran en el denominado Proyecto TENET (Teacher Education Network) cuya finalidad es sensibilizar al alumnado de magisterio para que a través de las materias implicadas en el currículum y su participación en las aulas durante las prácticas escolares, pueda llegar a configurar «un modelo educativo de género» que tienda a eliminar los sesgos y estereotipos sexistas en el aula. En estos programas se contemplan tanto las actitudes del alumnado, como los contenidos específicos y el empleo de una metodología adecuada. Son programas de *investigación-acción* que utilizan como recursos metodológicos la observación de la realidad (participante y no participante), registros de anécdotas, análisis de diarios, e informes de prácticas que permiten al alumnado tanto el aprendizaje de las destrezas relacionadas con la investigación educativa como la reflexión y crítica de las propias actuaciones dentro del aula. Estos programas TENET también se han desarrollado en otros países de la Comunidad Económica Europea (Palma, 1991; Arnesen, 1992; Wilson, 1991). También en el Congreso Internacional sobre Investigación, Docencia y Feminismo celebrado en Donostia en diciembre de 1992, se presentaron algunos programas para la formación de alumnos de magisterio en el ámbito no sexista y las evaluaciones realizadas (Nuño, 1992).

Por otro lado, la escasa representación de la mujer en las posiciones de liderazgo educativo en las escuelas primarias y secundarias, ha potenciado diferentes investigaciones feministas para tratar de identificar las razones de esta distribución desigual. Se ha examinado la estructura organizativa de las escuelas, las experiencias de las maestras, los obstáculos para la promoción de la mujer y la teoría y práctica del liderazgo educativo (Donn, 1986,1987; Strachan, 1991; Styfield, 1991; Thomson, 1988). Se implementan programas diseñados para realzar el desarrollo personal y profesional de las mujeres que aspiran a posiciones de liderazgo educativo (Strachan, 1993). La evaluación de estos programas se lleva a cabo a partir de la metodología de la investigación-acción. Una de las principales metas de la investigación en este ámbito es analizar el proceso y la dinámica del cambio en la mujer y en el sistema en el que ella trabaja. Se utilizan básicamente *entrevistas informales*.

3.4. Género y Educación para la carrera

La literatura vocacional en los años 90 continua explorando la existencia de

diferencias de género en la elección de carrera. Piel (1991), a partir de una amplia revisión bibliográfica de diferentes estudios e investigaciones en este ámbito, concluye que las diferencias tipificadas sexualmente en las elecciones de carrera son bastantes persistentes y no son fácilmente alteradas por la información ocupacional.

En este ámbito, la investigación sobre diferencias de género ha trabajado sobre *todo estudios de tipo descriptivo y correlacional*, empleando pocos recursos en *investigaciones experimentales o cuasi experimentales* sobre el efecto que pueden producir programas educativos en la evolución y cambio de estas diferencias (Bartolomé, 1991).

Algunas investigaciones desde los años 80, utilizando la metodología ex-post-facto, relacionan las diferencias de género con las expectativas de autoeficacia profesional en estudiantes universitarios. Estas investigaciones ponen de manifiesto la importancia de esta variable en el desarrollo profesional de la mujer. Entienden la autoeficacia como *los juicios individuales que el sujeto hace de su capacidad para organizar y ejecutar cursos de acción requeridos para alcanzar tipos de ejecuciones diseñadas* (Bandura, 1986). Así, Matsui, Ikeda y Ohnishi (1989), utilizando como sujetos a 78 hombres y 81 mujeres japonesas —todos ellos estudiantes universitarios de los primeros cursos— examinan la contribución de cuatro clases de socializaciones tipificadas sexualmente (idénticos modelos de rol de sexo en las profesiones, estereotipos del rol de sexo, percepciones de la propia masculinidad-feminidad y confianza en su ejecución en matemáticas) a las diferencias sexuales en autoeficacia, en 10 profesiones predominantemente masculinas y 10 predominantemente femeninas. Los hombres presentan una autoeficacia equivalente en los dos tipos de profesiones, mientras que las mujeres presentan una autoeficacia mayor en las profesiones predominantemente femeninas y menor en las masculinas.

Estos resultados confirman los hallazgos de Betz y Hackett (1981) a pesar de las diferencias entre los dos estudios en lo referente a las profesiones usadas y el contexto cultural y social de las muestras utilizadas (estudiantes japoneses y norteamericanos). Además, las mujeres presentaron una autoeficacia profesional más baja en las profesiones predominantemente masculinas que en las predominantemente femeninas, dado que creían tener menos modelos de rol femenino en las profesiones predominantemente masculinas, percibían la feminidad por ellas mismas y tenían una menor confianza en su ejecución en matemáticas.

Estas conclusiones establecen, conjuntamente con las de Betz y Hackett, fuertes patrones de diferencias de sexo en la autoeficacia profesional de alumnos universitarios.

Un estudio posterior de Stickel y Bonett (1991) analiza la influencia de la autoeficacia profesional para combinar una profesión con el hogar y la familia. Sus resultados vuelven a plantear la existencia de diferencias de género. Las mujeres creen tener mayor eficacia para combinar profesiones tradicionales con la familia y el hogar que los hombres. También creen poder fallar en profesiones no tradicionales no solo porque dudan de su propia habilidad para desarrollar tales profesiones sino también porque dudan de su habilidad para poder combinar sus requerimientos con las responsabilidades de la familia y el hogar, expectativa socialmente impues-

ta. El hombre puede reaccionar a estas presiones sociales o estar socializado de forma diferente, lo cual sugiere un nuevo foco de estudio para futuras investigaciones.

Diversas investigaciones han analizado las diferencias de sexo en la madurez vocacional (King, 1989) y los determinantes de la conducta vocacional de las mujeres utilizando complejos análisis estadísticos. Así, Fassinger (1990) mediante la metodología del *análisis causal*, sugiere que una combinación de habilidad, actitudes del rol de género y algunas características de personalidad, puede ayudar a explicar las elecciones de carreras no tradicionales de las mujeres jóvenes.

Farmer (1985) presenta un modelo multidimensional para determinar la influencia de diferentes variables (edad, género, etnicidad, habilidad verbal y matemática, soporte paterno, soporte del profesor y soporte al trabajo de la mujer) en la motivación del logro y motivación para la carrera de jóvenes de ambos sexos (N = 1.863). Mediante el programa Lisrel IV identifica cuatro modelos causales diferentes: uno para el *nivel de aspiración* para la carrera (nivel de la profesión escogida), otro para el *dominio* (motivación para lograr tareas), otro para el *compromiso hacia la carrera* del chico y otro para el *compromiso hacia la carrera* de la chica.

El estudio de Tusin (1991) investiga las relaciones del autoconcepto social y académico con la elección de la enseñanza como carrera, en el contexto universitario. Desarrolla un *modelo causal* en el que relaciona cada una de las tres medidas dependientes (la elección de la enseñanza como carrera, el autoconcepto social y el académico, después de 9 años de haber iniciado los estudios universitarios) con características de «background», elección de carrera y el autoconcepto académico y social a su entrada a la universidad, factores de la institución postsecundaria, experiencias postsecundarias y la primera ocupación como profesor. La muestra utilizada en su estudio está formada por 2.441 mujeres. Los datos longitudinales se recogieron en 1971 y 1979-80. Se encontraron efectos significativos diferentes entre los profesores de enseñanza primaria y secundaria.

Como señalábamos antes, hay pocos estudios en este ámbito que hayan utilizado la metodología *experimental o cuasiexperimental*. En la revisión realizada sólo hemos encontrado el de Yanico (1978) sobre los sesgos debidos al sexo en la información profesional de alumnos universitarios. Esta autora explora los efectos en las actitudes de los sujetos hacia diferentes tipos de profesiones cuando se introducen cambios en la información ocupacional. Estas informaciones sobre las profesiones eran de tres tipos: sesgadas sexualmente (tanto el contenido como el lenguaje utilizado), sin sesgos implícitos (sin sesgos en el contenido presentado) y sin sesgos explícitos (tanto el contenido como el lenguaje utilizado en la descripción estaban exentos de sesgos debidos al sexo). El objetivo era estudiar si la neutralización o no del lenguaje en las descripciones sesgadas puede llevar a actitudes sesgadas entre los sujetos respecto a la adecuación de las profesiones para hombres y mujeres. Los resultados mostraron que: a) los sujetos inicialmente valoraron la mayoría de las profesiones en una dirección estereotipada; b) en profesiones tradicionalmente masculinas, después de una información sesgada, tienden a incremen-

tar el sesgo; c) los sujetos estuvieron menos dispuestos a cambiar su actitud sobre la adecuación de ocupaciones «femeninas» para el hombre que sobre ocupaciones «masculinas» para la mujer; y d) el material sin sesgo explícito no producía diferencias en las actitudes de los sujetos.

Por otro lado, se han *diseñado, implementado y evaluado* toda una serie de programas de intervención para ayudar a las adolescentes a ensanchar sus elecciones vocacionales, especialmente en lo relativo a las elecciones de carreras no tradicionales (Cini y Baker, 1987; Cramer, Wise y Colburn, 1977; Johnson y Externa, 1982; Prediger y Noeth, 1979; Wilson y Daniel, 1981 y Reat-Poteat y Martin, 1991). Estos estudios sugieren que la educación de la mujer sobre carreras no tradicionales se logra mejor con una «aproximación integrada» que incluya tanto ejercicios de desarrollo del autoconcepto como actividades de exploración de la carrera.

La evaluación de la efectividad a corto plazo de estos programas, casi siempre realizada a partir de *entrevistas o cuestionarios* aplicados a las participantes, señala el éxito de ellos en: a) el cambio de actitudes hacia el trabajo; b) la reducción de los estereotipos del rol del sexo; c) el incremento del conocimiento de las profesiones y el conocimiento de la mujer del mundo del trabajo; d) el logro de un mayor autoconcepto vocacional y e) el incremento del interés hacia trabajos no tradicionales (Guttman, 1991). No obstante, faltan estudios que analicen los efectos a largo plazo de este tipo de programas.

En nuestro contexto, se han elaborado una serie de programas dirigidos específicamente desde gabinetes de orientación como es el caso de la mayoría de experiencias que se presentaron en las «Primeras Jornadas de intercambio de experiencias de orientación para la igualdad de oportunidades entre ambos sexos», celebradas en Madrid en noviembre de 1990 o los «Talleres de Orientación» realizados en diferentes centros escolares, tanto de E.G.B. como de secundaria, en L'Hospitalet de Llobregat, durante los cursos 1990-91 y 1991-92, dentro del Plan de Acción para la Igualdad de Oportunidades de la Mujeres, promovido por el ayuntamiento de esta ciudad (Escolà y otros, 1991; López y Sobré, 1993). Estos programas de orientación no sexista incluyen estrategias de actuación educativa, materiales y recursos a utilizar (guías de autorreflexión, documentación, cuentos o relatos que intentan romper estereotipos, etc.).

Cuando se lleva a cabo la evaluación de estos programas, se realiza a partir de *cuestionarios* aplicados a los alumnos participantes, sus profesores y padres, una vez finalizada la experiencia; *análisis de grabaciones de las diferentes reuniones* con el profesorado y de los diarios del alumnado y profesorado. En otros casos se llevan a cabo *observaciones* de las actividades realizadas dentro del programa y *entrevistas* a los participantes en el mismo.

La literatura vocacional de estos últimos años apunta dos importantes áreas en las que el estudio e investigación sobre género y desarrollo vocacional puede progresar.

Por un lado, aparece en publicaciones recientes la necesidad de estudiar a fondo

la interacción del género con los factores de raza y etnicidad, interacción que no ha sido trabajada con suficiente rigurosidad (Scher y Good, 1990). Reid y Comas-Díaz (1990) señalan además que la literatura profesional no debe olvidar la interacción raza-género, omitiendo tópicos tan importantes como la identificación dual y la doble discriminación. Una serie de artículos recogidos en el volumen 22, nº 7/8 (1991) de la revista «Sex Roles» discuten toda una serie de tópicos relevantes para entender a la mujer de color, específicamente en lo que hace referencia a mujer y trabajo. Gilkes (1990) analiza el trabajo de mujeres negras americanas y mujeres blancas, señalando que los estereotipos asociados a la mujer negra respecto a los roles de su trabajo, son extremadamente persistentes. Utiliza para su estudio *entrevistas en profundidad*. Burlew y Johnson (1992) investigaron las diferencias en las experiencias vocacionales de mujeres negras que estaban trabajando en carreras no tradicionales (medicina, derecho e ingeniería) y tradicionales (trabajo social, enseñanza y counseling). El objetivo de su investigación fue estudiar las barreras y oportunidades de estas mujeres cuando trabajan en profesiones no tradicionales y las tensiones provenientes del mundo familiar o profesional que incidían sobre ellas. Esta investigación *ex-post-facto* confirma los resultados de otras investigaciones similares al concluir que las mujeres negras que ejercen profesiones no tradicionales tienen más barreras para el éxito profesional, tales como discriminación racial y de género, oportunidades limitadas para el desarrollo político y dudas de sus colegas sobre su competencia profesional, además de un menor apoyo de sus iguales. Las mujeres casadas indicaron tener mayores barreras sociales en las profesiones no tradicionales que en las tradicionales y un mayor desacuerdo marital respecto a la carrera. Estos hallazgos refuerzan la necesidad de futuras investigaciones sobre la mujer en carreras no tradicionales, tanto por las barreras adicionales que conllevan como por la tensión incrementada que supone para la mujer el combinar las responsabilidades familiares con las profesionales, utilizando diseños multivariados más potentes y réplicas en diferentes situaciones y contextos multiculturales.

Una segunda nueva área de trabajo nos lleva explícitamente al tema del «trabajo dentro de la vida». El sistema de relaciones tiene tradicionalmente una importancia crucial en las vidas de las mujeres. Consecuentemente, alguna literatura actual insinúa que se necesita describir explícitamente el trabajo de la mujer en términos de relaciones más que como un compartimento estanco que es como han sido tratadas tradicionalmente las carreras del hombre (Piel, 1993). Los modelos vocacionales de las mujeres representan complejos esfuerzos por equilibrar sus vidas, modelos que son diferentes pero no inferiores a los modelos vocacionales típicamente masculinos (Gallos, 1989). Estudios recientes confirman que las diferencias de género son el resultado de un proceso de negociación con los otros a lo largo de la vida. Esta perspectiva «negociadora» en el estudio del género arroja una nueva luz sobre las diferencias familiares en la planificación de la vida y la carrera de mujeres y hombres (Espín y Figuera, 1991; Piel, 1993). Mediante el proceso de socialización, el hombre y la mujer tienden a desarrollar diferentes orientaciones con respecto a sus relaciones interpersonales y profesionales. Estas diferencias

interactúan con las normas socioculturales, produciendo diferentes oportunidades y demandas de estilo de vida.

3.5. Hacia una investigación integrada

Como ya hemos apuntado en el apartado anterior, algunas investigaciones ponen de manifiesto la necesidad de estudiar a fondo la interacción del género con los factores de raza, etnicidad y clase social. La mayoría de la literatura sobre género en educación ha ignorado las vías por las que las diferencias basadas en la clase social, origen regional, etnia o raza conllevan diferencias de género, con el resultado de que sólo emerge una pintura parcial de la mujer en el sistema educativo (Jiménez Fernández, 1990; Wilson, 1991). La investigación sobre desigualdad de oportunidades educativas en términos de clase social o etnia se ve empobrecida por una falta de atención a las diferencias de género. Una razón para tal deficiencia en la investigación puede estar ligada a la gran complejidad de la interrelación entre los factores envueltos.

Algunos pasos se han dado en este sentido. Así, el análisis teórico de Gewirtz (1991) sobre sexismo y racismo en educación, explora las aproximaciones teóricas predominantes en cada uno de estos dos campos y la dirección seguida en las estrategias que se han adoptado para el cambio. También Haw (1991) analiza la literatura reciente sobre las interrelaciones entre género y raza y la problemática que puede suponer al profesorado no tener en cuenta estas interacciones y no emplear en sus prácticas escolares, estrategias que aborden conjuntamente las cuestiones sobre etnia y género.

La investigación de Reay (1991) examina la intersección entre el género, la raza y la clase social en la escuela primaria. A partir de las grabaciones en vídeo de dos grupos de niños (un grupo mixto y otro sólo de niñas) cuando están trabajando en un mismo tipo de tarea, analiza sus discusiones de grupo. El código de análisis comprende las siguientes categorías: conductas lingüísticas, dependencia del profesor, modelos de trabajo individual o en grupo, división del trabajo y modelos de trabajo jerárquico o cooperativo. Los resultados de su estudio apoyan el punto de vista de que el impacto del género en las niñas no es uniforme ni consistente cuando se introducen las variables de raza y clase social. Lo cual lleva a esta autora a afirmar que estas tres variables están muy interrelacionadas. Otras investigaciones han relacionado el género y la etnicidad al estudiar los efectos de estas variables sobre las percepciones de los alumnos en el ámbito de las matemáticas o el lenguaje (Birenbaum y Kraemer, 1992).

Se necesitan más investigaciones que analicen las intersecciones del género con la raza y clase social para poder ayudar al profesorado a abordar sus prácticas cotidianas.

Por otra parte, Robinson (1993) presenta un análisis interesante donde explora las intersecciones múltiples y dinámicas entre género, raza, cultura y clase social, y su influencia sobre la formación de la identidad psicosocial, ilustrándolo a partir de un *estudio de casos*.

Como apunta Bartolomé (1991) las recientes publicaciones sobre educación intercultural o educación para la justicia, han puesto de relieve que en los estudios sobre género y educación, se ha olvidado o pasado por alto, a menudo, las variables etnia y clase social, como variables moduladoras de las diferencias de género halladas en la sociedad. Esta advertencia y la consideración de los límites con que tropezamos al evaluar generalmente los programas de educación no sexista, han conducido a planteamientos más integrados en los que se conecta la educación para la paz (Feminario de Alicante, 1987; Brotons, 1987) o la educación global (Leming, 1992), por ejemplo, con una auténtica educación no sexista y no racista incluyendo todas en el seno de una *educación para la responsabilidad social*.

4. MODELOS DE INVESTIGACIÓN EN LA INTERVENCIÓN EDUCATIVA ASOCIADA A LAS DIFERENCIAS INDIVIDUALES

4. 1. Introducción

No parece existir una correspondencia entre la larga tradición del estudio de las diferencias individuales en psicología y pedagogía, con un trabajo similar de investigación sobre la intervención basada en las diferencias.

La revisión bibliográfica realizada para el presente trabajo desde mediados de los 80 hasta la actualidad nos muestra en el campo de la educación diferencial un claro predominio de investigaciones de tipo descriptivo estudiando las relaciones entre las aptitudes o actitudes y el rendimiento, sobre aquellas otras que tienen como objetivo la intervención asociada a las diferencias. Dentro del marco de las Ciencias de la Educación el objetivo principal debería ser el de buscar principios, métodos o programas pedagógicos de acción sustentados empíricamente y que favorezcan una verdadera educación diferencial. Sin embargo parece existir un mayor interés por parte de los investigadores en efectuar estudios descriptivos sobre las diferencias, subsidiarios en su mayoría de las conceptualizaciones realizadas por psicólogos cognitivos y del aprendizaje.

Una gran parte de las investigaciones sobre las diferencias son aplicaciones al medio escolar de los constructos operacionalizados por la psicología cognitiva. Se constata entre los objetivos de la mayoría de los trabajos la búsqueda de la interacción entre la modalidad de tratamiento educativo empleado y las características de los alumnos, siguiendo la corriente que apuntaba Tejedor en 1988. Entre los aspectos más trabajados destacan: la relación entre estilos cognitivos, motivación, estilos de aprendizaje y rendimiento.

4.2. Investigación sobre intervenciones basadas en los Estilos Cognitivos

Las investigaciones sobre estilos cognitivos han proliferado en el marco educativo desde las formulaciones iniciales de este constructo en 1962 por Witkin y en 1963 por Kagan, Moss y Sigel. En un principio los trabajos empíricos se centraron

en la conceptualización y medida del constructo (Carretero, 1982 y Palacios, 1982).

En la actualidad los trabajos siguen fundamentalmente la conceptualización de Kagan (1981, 306) que los define como «variación individual de los modos de percibir, recordar o pensar; o como formas distintas de aprender, almacenar, transformar y emplear la información». Esta definición aproxima el constructo a lo que otros autores denomina estilos de aprendizaje, o estrategias de aprendizaje. Se considera, sin embargo, que el estilo cognitivo es una característica fija del individuo, mientras que los estilos y estrategias de aprendizaje pueden variar según la situación, la tarea o incluso la edad del sujeto (Riding y Sadler-Smith, 1992).

Las investigaciones realizadas en relación con las diferencias en los estilos cognitivos y los medios de enseñanza se encaminan a identificar los atributos fundamentales de los medios en relación con los efectos que éstos pueden tener en los procesos cognitivos de los sujetos, bien para determinar qué atributos específicos son más efectivos para determinados sujetos, bien para saber qué tipo de procesos cognitivos se implican en el procesamiento de materiales, codificados y presentados en forma diferente (Cabero y Salas, 1990; Riding y Sadler, 1992; Burwell, 1991; Post, 1987; Miller y Emiwovich, 1988), sobre todo medios informáticos y audiovisuales. Predomina la *metodología experimental* y los *diseños ATI*.

Los trabajos de investigación, siguiendo estas metodologías, se centran en dos tipos de cuestiones, por una parte ver si los sujetos con diferentes estilos cognitivos aprenden y rinden de forma diferente según el nivel de estructuración de los materiales de enseñanza, o bien ver si las tareas, como por ejemplo la programación en Logo, modifican el estilo cognitivo (Bradley, 1985; Clement y Gullo, 1984; Foux y Routh, 1984). Los trabajos efectuados para probar la relación entre la estructuración de los materiales y los estilos cognitivos se han enfocado desde dos vertientes, la audiovisual y la informática.

Los resultados obtenidos con medios informáticos destacan el aumento de puntuaciones en reflexividad de los sujetos que trabajan con cursos de programación Logo (Miller y Emiwovich, 1988). Se observa mayor efectividad en el aprendizaje si se empareja adecuadamente el estilo cognitivo con el formato de instrucción (Rowland, 1988).

En la primera destacan la investigaciones de Cabero (1989) y Cabero y Salas (1990), con una *metodología experimental*, que utilizan la versión clásica de los estilos cognitivos Dependencia /Independencia de campo y la Reflexividad /Impulsividad. Los resultados muestran que ambos estilos influyen en el rendimiento a corto y medio plazo a la hora de movilizar en el sujeto determinadas estrategias de procesamiento de la información y de organizar conceptualmente los estímulos procedentes del medio exterior. En un trabajo posterior Cabero y Salas (1990) concluyen que el estilo cognitivo del alumno funciona como elemento filtrador de los posibles efectos de los medios de enseñanza y por tanto influye en los productos finales de los alumnos. Faltan sin embargo trabajos que expliquen o describan los mecanismos de percepción, almacenamiento y procesamiento que tienen lugar en el sujeto en relación con diferentes medios y tareas de enseñanza.

Desde la vertiente informática, Riding y Sadler-Smith (1992) después de revisar los trabajos que identifican las semejanzas y diferencias entre las dimensiones medidas por diferentes estilos cognitivos proponen dos estilos que caracterizan las diferencias individuales: el Holístico-Analítico y el Verbal-Imaginativo. Según estos autores cada persona se ubica dentro del continuum en cada una de las dimensiones de estos dos estilos. Definen el estilo Holístico-Analítico como el tendente a procesar la información en conjunto o por partes; y el estilo Verbal-Imaginativo como el tendente a la representación mental de la información de modo lingüístico o por imágenes. Su investigación, con *diseño experimental*, se centra en estudiar los efectos en el aprendizaje de los dos estilos cognitivos según el grado de estructuración de los materiales de enseñanza, en sujetos de 14 a 19 años. Los resultados indican que los alumnos procesan la información de modo diverso según su estilo cognitivo, por tanto la estructura del material de aprendizaje afecta de forma distinta a las personas según su estilo cognitivo (la estructuración del material está relacionada con la dimensión Holístico-Analítica, y el modo de presentación y el contenido están relacionadas con la dimensión Verbal-Imaginativa). Para algunos estilos, como por ejemplo el Analítico-Imaginativo hay una gran diferencia entre las puntuaciones que se obtienen en cuanto a la efectividad del aprendizaje según se trabaje con una u otra versión de tratamiento. Se comprobó los alumnos dentro de los grupos Analítico-Imaginativo u Holístico-Verbal desarrollan estrategias alternativas compensatorias para las tareas de aprendizaje cuando éstas requieren un procesamiento de la información que no va de acuerdo con el propio estilo. Las personas caracterizadas por los estilos Holístico-Imaginativo o Analítico-Verbal tienen más dificultades para desarrollar estrategias alternativas.

Una línea de investigación prometedora es aquella que engloba diferentes trabajos destinados a probar la efectividad de los vídeos interactivos en la enseñanza, en cuanto a la reducción del tiempo de aprendizaje y la mejora en la retención y ejecución de lo aprendido. Hannafin, Phillips y Tripp (1986) investigan el papel del control del aprendizaje por parte del alumno, considerado como un método de enseñanza para tratar las diferencias individuales, trabajando con videodisco interactivo. Burwell (1991) trata de establecer la interacción entre los estilos cognitivos Dependiente e Independiente de campo y el control del aprendizaje en la enseñanza asistida por ordenador utilizando el videodisco interactivo. Emplea un *diseño ATI* con tres tipos de métodos de enseñanza con diferente grado de control del aprendizaje por parte del alumno. Los resultados del trabajo muestran que hay una relación significativa entre los estilos cognitivos y los tratamientos como se puede apreciar en la figura 5.

Hay una interacción «disordinal» como muestra la intersección entre los vectores de las dos variables. Los Independientes de Campo funcionan mejor con un método de enseñanza que tiene estructurada la secuencia de aprendizaje, mientras que los Dependientes de Campo funcionan mejor con un método instructivo que les permita el control del propio aprendizaje.

Se calculó también la interacción entre el estilo cognitivo y los tratamientos

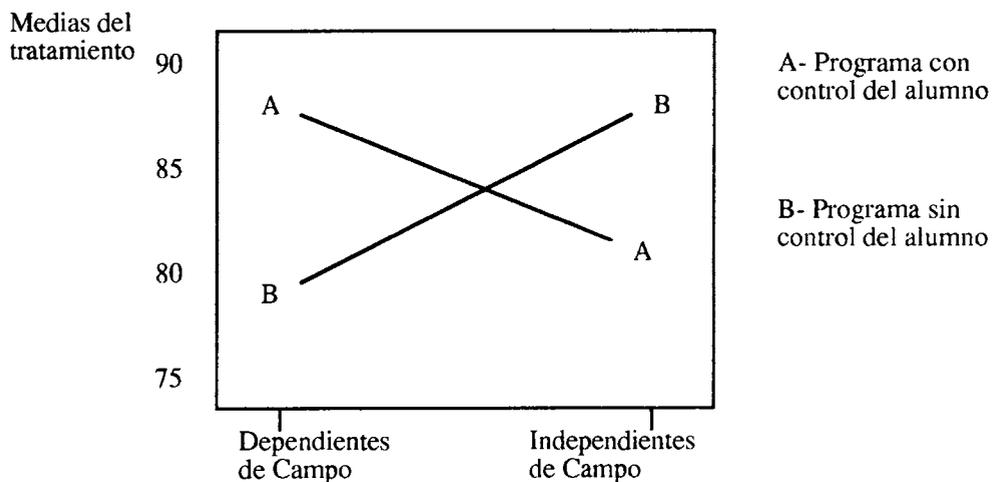


Figura 5
Interacciones Estilo Cognitivo-Tratamiento. Puntuaciones del psttest.

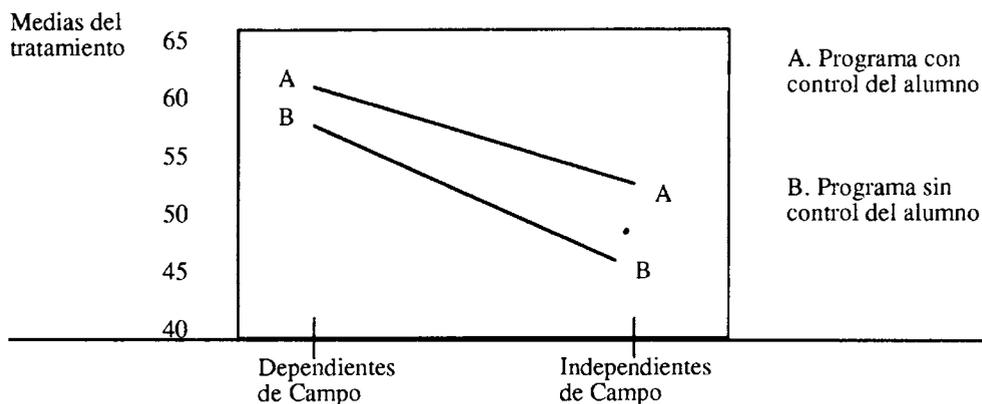


Figura 6
Interacción-Aptitud-Tratamiento. Tiempo de estudio.

instructivos en cuanto al tiempo invertido en el estudio. Los resultados se muestran en la figura 6. El gráfico muestra una interacción ordinal ya que no hay intersección entre los vectores de los dos tratamientos, pero tampoco corren paralelos. Se constata que los Dependientes de Campo tardan más que los Independientes de Campo en finalizar o completar el estudio de las lecciones propuestas.

Otro de los campos a los que se ha aplicado la investigación sobre los estilos cognitivos es el de la modificación de conducta. Partiendo de los trabajos realizados en los años 70 en los que se ponía de manifiesto la mayor impulsividad y dependen-

cia de campo de los niños hiperactivos sobre los normales tanto en preescolar como en cursos más avanzados, Orjales y Polaino-Lorente (1992), con un *diseño cuasiexperimental*, tratan de comprobar el grado de modificabilidad de la conducta con dos tipos de metodologías, una de tipo cognitivo-conductual y otra de tipo tradicional de refuerzo de los aprendizajes básicos. Los resultados de la investigación confirman la presencia de un Estilo Cognitivo Dependiente de campo para los niños hiperactivos, pero no confirman un Estilo de Impulsividad aunque hay una cierta tendencia a ella. El grupo que sigue la metodología cognitivo-conductual evoluciona hacia la rapidez exactitud, mientras que el grupo con método tradicional avanza hacia la lentitud exactitud.

4.3. Investigaciones sobre la intervención en Estilos de Aprendizaje

Los estilos de aprendizaje son conceptualizados de forma distinta por los diferentes autores, unos los caracterizan como el modo preferente de aprender escogido por los alumnos (Dunn y Dunn, 1975), para otros se relacionan con el uso de un determinado tipo de estrategias para el estudio por parte de los alumnos (Selmes, 1987; Schmeck, 1988). El interés de los investigadores se centra fundamentalmente en identificar los estilos de aprendizaje para diseñar métodos de enseñanza adecuados a las características diferenciales de los alumnos.

La línea de investigación en la que se inscriben los trabajos de Dunn y Dunn (1975) y Dunn, Dunn y Price (1977), se caracteriza por elaborar instrumentos de medida para examinar las preferencias de los alumnos por distintos modos de aprender, teniendo en cuenta variables ambientales, individuales y sociales. Estas autoras analizan las diferentes metodologías de enseñanza proponiendo las más adecuadas para cada tipo de preferencias (Dunn y Dunn, 1984), a partir de los resultados obtenidos en las investigaciones sobre las preferencias en función del sexo y la edad, y su evolución a lo largo de la escolaridad, así como la relación entre autoconcepto y estilo de aprendizaje, y rendimiento en lectura y matemáticas. Faltan sin embargo investigaciones para comprobar la efectividad de estas metodologías.

Los trabajos de Selmes y Schmeck, ponen en relación las metas de aprendizaje de los alumnos con las estrategias empleadas en el estudio. El primero sigue una *metodología cualitativa*, empleando la entrevista, y el segundo con *metodología cuantitativa*, analizando los resultados de cuestionarios. Se caracterizan dos enfoques, superficial y profundo (Selmes, 1986), y en este último Schmeck (1988) distingue dos tipos de sujetos los de enfoque conceptualizador y los de enfoque personalizador. Los resultados de ambos autores son coincidentes en la mayor efectividad del enfoque profundo sobre el superficial. En esta línea un trabajo interesante es el de McRobbie (1991), que utilizan un *diseño cuasiexperimental* para detectar si había una relación entre elementos específicos del estilo de aprendizaje del alumno (preferencia por un tipo u otro de estrategias de aprendizaje), la organización cognitiva (modo de procesar la información: factual o conceptual) y la

metodología de enseñanza. Los resultados de la investigación demuestran que la preferencia cognitiva del estudiante, definida por las estrategias de aprendizaje empleadas, fue un predictor significativo tanto para la organización cognitiva como para los resultados del aprendizaje.

4.4. Investigaciones para el desarrollo de las habilidades intelectuales

El enfoque de la psicología cognitiva que estudia las operaciones y los procesos mentales subyacentes a la resolución de tareas y problemas, ha generado un gran número de investigaciones para mejorar las habilidades cognitivas y las estrategias de aprendizaje de los sujetos, sin contar aquellas que las emplean como variables diferenciales para ver la eficacia de diferentes metodologías instructivas o estructuración de situaciones de aprendizaje.

Según Glaser (1984) y Nickerson y otros (1987), las líneas de intervención cognitiva se pueden clasificar en programas orientados al desarrollo de: procesos (Feuerstein, 1980; Proyecto Inteligencia, 1983; PAR, 1986); estrategias mentales, o heurísticos como el De Bono (1983); pensamiento formal de tipo piagetiano (Fuller, 1980); el desarrollo del lenguaje (Scardamalia 1979), y estrategias metacognitivas (Campiono, 1982; Lipman, 1980).

Si bien se reconoce que la mayor parte de los estudios de entrenamiento en estrategias cognitivas y metacognitivas muestran que el aprendizaje y la memoria pueden desarrollarse a través de la enseñanza (Corno y Snow, 1986), y que se puede aprender a transferir desde unas situaciones estratégicas a otras, sobre todo si se entrena en tareas de tipo cognitivo generalizable (Dansereau, 1985). En la actualidad el interés de los investigadores se ha ido desplazando hacia la integración de la enseñanza de los procesos cognitivos con los contenidos del currículum (Monereo, 1993) pues se considera que el aprendizaje de estrategias no puede separarse del soporte informativo que manejan y, además, este soporte posee una estructura lógica singular que afecta a las estrategias dándoles una cualidad de específicas, es decir la materia o asignatura media-tiza las estrategias (Brandsford y otros 1990; Nisbet, 1991).

El interés de los investigadores por la enseñanza de estrategias dentro de contextos concretos de aprendizaje, es decir dentro de cada una de las materias del currículum se ha visto estimulado por los trabajos de Glaser (1984) en los que demuestra la importancia del conocimiento previo en la resolución de problemas, o los de Mayer (1987) sobre la resolución de problemas matemáticos, junto con los resultados de la evaluación de los distintos programas de entrenamiento en operaciones cognitivas básicas (en los que se concluye que los sujetos mejoran sus habilidades cognitivas pero luego tienen dificultades para trasladarlas al aprendizaje de materias específicas). Sin embargo este enfoque de integración en el currículum está teniendo dificultades para ser trasladado al aula al tropezar con actitudes de resistencia y rechazo por parte del profesorado (Pressley y otros, 1990; Monereo, 1993).

Los enfoques actuales sobre la intervención en las estrategias de aprendizaje intentan superar la división existente entre materia-profesor y alumno y consideran

que en los contextos académicos «las estrategias de aprendizaje son procesos estructuralmente compartidos por el alumno, moduladas por los contenidos que se aprenden y por cómo son enseñadas por el profesor» (Bernad, 1993,16).

La *investigación cooperativa* se comienza a utilizar como un medio de introducir programas de entrenamiento en estrategias de aprendizaje en las distintas materias del currículum a través de los contenidos y actividades que se realizan en el aula a través del currículum (Palincsar y otros, 1988; Duffy y Roehler, 1989). Estos autores defienden que es necesaria una estrecha colaboración entre investigadores y formadores para garantizar el mantenimiento y la eficacia de las estrategias entrenadas. Monereo (1993) siguiendo un modelo de *investigación-acción* plantea un estudio que incluye una formación del profesorado en el aprendizaje y enseñanza de estrategias para comprobar posteriormente si hay diferencia en la actuación de los alumnos comparando la actuación de los profesores formados y no formados, en una misma escuela. Los resultados de la primera fase muestran una mejora significativa de los profesores que han recibido una formación tanto en conocimientos estratégicos como en el uso personal de estas estrategias para aprender un texto determinado, si bien no existen diferencias entre los profesores formados y el resto en cuanto a su actuación docente.

La mayoría de las investigaciones sobre estrategias en la actualidad se vinculan al aprendizaje de materias específicas. La metodología que predomina en estos trabajos es la *empírico analítica*. Entre los tópicos más trabajados podemos citar los relacionadas con:

— el aprendizaje de procedimientos, con alumnos de EGB, BUP y universitarios (Pozo y otros, 1993). En un primer estudio analizan las actividades de enseñanza/aprendizaje llevadas a cabo en una situación real de aula empleando la *metodología observacional*, para ver los mecanismos de influencia docente sobre la adquisición de procedimientos de representación gráfica de información en dos asignaturas diferentes. En una segunda investigación *ex post facto* estudian los procesos de codificación implicados en la lectura de mapas geográficos, utilizando una metodología de enseñanza experimental diseñada para comprobar la eficacia de distintos tipos de instrucción procedimental en el aprendizaje de los diferentes tipos de información incluidos en un mapa.

— el desarrollo de la comprensión lectora (Alonso Tapia y Mateos, 1985; Vidal-Abarca y Gilabert, 1989 y 1991; Carriedo, 1993; Solé, 1993). Utilizan *diseños cuasiexperimentales* de grupos naturales equivalentes con pretest y doble posttest (final y de seguimiento).

— la composición escrita (Bereiter, 1980; Bereiter y Scardamalia, 1987). Valoran su programa de entrenamiento en la composición escrita, combinando el uso de ayudas externas con el modelado en la forma de pensar y el entrenamiento en el uso directo de las estrategias, mediante *diseños cuasiexperimentales* con grupos naturales equivalentes.

— estrategias motivacionales en la instrucción (Ames 1992; Blumenfeld, 1992; Alonso Tapia, 1993). Este último elabora un cuestionario para ver las diferencias

entre los profesores en cuanto a sus comportamientos motivacionales hacia el aprendizaje, como base previa para un programa de formación de profesores en el conocimiento y aplicación de estrategias motivacionales durante su actividad docente.

— la mejora de las habilidades para el estudio (Selmes, 1986, 1988). En el primer trabajo con una metodología cualitativa basada en el *análisis de contenido de entrevistas estructuradas*, diferencia dos tipos de enfoques, superficial y profundo. Posteriormente aplica un programa para la mejora de habilidades para el estudio que incluye tareas de modelado en el profesor y estrategias a seguir por el alumno. Valora los resultados mediante entrevistas en las que se investiga la mejora de la confianza personal, la competencia y habilidades para el estudio.

4.5. Investigaciones sobre intervención en la estructuración de las situaciones de aprendizaje

Desde los años 70 existe un debate considerable sobre cuál es el modo más deseable de estructurar las situaciones de aprendizaje para promover el logro en los alumnos. Se han llevado a cabo numerosos estudios sobre las diferencias individuales en la comprensión, retención o calidad de estrategias de razonamiento en función de la estructuración de las situaciones de aprendizaje. Johnson y Johnson (1990) en un *metaanálisis* sobre 323 investigaciones en las que se compara el impacto relativo de situaciones de aprendizaje cooperativas, competitivas o individualísticas en el logro de los alumnos, concluyen que se da un mayor logro en situaciones cooperativas que en las competitivas o individualísticas. El aprendizaje cooperativo favorece que los alumnos empleen más a menudo estrategias de razonamiento complejo que cuando están en situación individual o competitiva.

El tema clave de estas investigaciones está, por una parte en que no hay una unanimidad en cuanto a las medidas de logro empleadas, que van desde la comprensión y retención del material aprendido, a las calidad de las estrategias de razonamiento, a la generación de nuevas ideas y la transferencia de lo aprendido a otras situaciones; y por otra parte en especificar las variables que median en la relación entre aprendizaje cooperativo y logro.

A menudo se presta a confusión la identificación entre aprendizaje en grupo y aprendizaje cooperativo (Barnett, 1992), para que se pueda hablar de este último Slavin (1990) señala las condiciones que ha de cumplir, entre las que destacan, la heterogeneidad de los componentes, las metas positivas como equipo, la interdependencia entre los participantes, el desarrollo de habilidades de comunicación entre los miembros, la responsabilidad individual en el aprendizaje, y la evaluación del proceso del trabajo en equipo.

Algunos trabajos de *investigación evaluativa* como los de King (1990) se centran en desarrollar la efectividad de algunas de las estrategias propuestas por el aprendizaje cooperativo, como por ejemplo las habilidades de comunicación, diseñando estrategias de cuestionamiento por pares y entrenando a los sujetos en ellas para facilitar el aprendizaje de materiales trabajados en clases de lectura. Como

conclusión de esta investigación se destaca la importancia de entrenar y enseñar a los alumnos el proceso de formulación de cuestiones. Ya que los resultados indican que los estudiantes que participan en los grupos de discusión entrenados en el planteamiento de cuestiones responden más y mejor a las preguntas referidas a pensamiento crítico, dan más explicaciones y con respuestas más elaboradas y obtienen mejor rendimiento que los estudiantes del grupo de discusión que no ha sido entrenado.

Una vez identificadas las variables que caracterizan el aprendizaje cooperativo, y demostrada su superioridad sobre otras modalidades, el interés de los investigadores se ha diversificado. Una línea de trabajo se ha centrado en detectar los efectos de este tipo de estructuración de la situación de aprendizaje en variables de tipo actitudinal y otra, en variables de tipo cognitivo.

En el primer caso destacamos los trabajos desarrollados en España por Calvo (1990 y 1991), sobre la modificación de actitudes autoritarias y de prejuicios en alumnos de BUP, siguiendo un modelo de *investigación en el aula*. Los resultados ponen de manifiesto después de la experiencia una reducción de las actitudes autoritarias y dogmáticas de los alumnos y una mejora del clima de clase en el grupo que participó en la experiencia con respecto al grupo de control. Se detectó también una mejora en el rendimiento y la motivación de los alumnos.

En el segundo están los trabajos de Brown y Campione (1986) en los que demuestran que el aprendizaje cooperativo es un componente para el aprendizaje cognitivo y los de Slavin y Stevens de la Universidad John Hopkins, siguiendo un *modelo empírico analítico*, que han desarrollado diversos programas de aprendizaje cooperativo para la enseñanza de la lectura y la composición escrita. Los resultados de tres *experimentos de campo* comparando el programa cooperativo de la enseñanza de la lectura y la escritura en los grados elementales, con los métodos tradicionales muestra que los alumnos que aprenden con el programa tienen mejores resultados en los tests standarizados (Stevens y otros, 1987; Stevens, Slavin y Farnish, 1989). En la actualidad este grupo de investigadores está interesado por el estudio de la combinación de los métodos de la instrucción directa en las estrategias de comprensión lectora y el aprendizaje cooperativo. En una investigación *ex post facto* realizada por Stevens, Slavin y Farnish en 1991 en 30 escuelas elementales para ver el impacto de la instrucción directa en las estrategias de comprensión lectora y el grado en que el aprendizaje cooperativo potencia la efectividad en el uso de estas estrategias por parte de los alumnos, los resultados mostraron que el proceso de aprendizaje cooperativo integrado con la instrucción directa es el procedimiento de enseñanza más efectivo.

La generalización progresiva del ordenador como medio de enseñanza en los centros ha suscitado de nuevo el interés por los métodos de enseñanza individualizada al contar con un medio potente para desarrollar diversos modos, secuencias y estructuración de las tareas (Carrier y Jonassen, 1988). Sin embargo el costo excesivo que este tipo de propuestas podría suponer junto con las ventajas que se derivan para el logro del alumno de la situación de aprendizaje cooperativo ha derivado el

interés de los investigadores por estudiar la introducción de la enseñanza asistida por ordenador dentro de un contexto de aprendizaje cooperativo. Los primeros trabajos han tenido como objeto clarificar el tipo de estrategias de enseñanza por ordenador que pueden ser aplicadas a un aprendizaje cooperativo, ya que no se pueden trasladar sin más a ese contexto las propias de la enseñanza individualizada. Watson (1990-91) y Hooper (1992) presentan un resumen de las investigaciones realizadas por el mismo y otros autores para clarificar los efectos del aprendizaje cooperativo empleando el ordenador y poner así las bases para el diseño de software para grupos.

A partir de los trabajos de Resnick (1987) en los que se demuestra las diferencias entre el tipo de aprendizaje requerido en el medio escolar y fuera de él, O'Donnell Dansereau y Rocklin (1991), con un *diseño cuasiexperimental* plantean una investigación sobre aprendizaje cooperativo en contextos no académicos. El objetivo de la investigación es ver los efectos de las diferencias individuales en el aprendizaje, ejecución y retención de procedimientos concretos dentro de un contexto variado de estrategias de aprendizaje cooperativo. Como variables diferenciales se consideran el nivel de dominio de vocabulario y el estilo cognitivo. Los resultados muestran diferencias en la ejecución y recuerdo de procedimientos entre los dependientes e independientes de campo según el tipo de estrategias de aprendizaje cooperativo con las que se ha trabajado; así como diferencias a favor de los sujetos con mejor nivel de vocabulario, independientemente del tipo de estrategia empleada. Los autores concluyen señalando la importancia de considerar los diferentes métodos de aprendizaje en relación con las diferencias de estilo cognitivo de los alumnos.

4.6. Investigaciones sobre el desarrollo de la motivación académica

Los trabajos empíricos en torno a la motivación académica han experimentado un creciente revitalización a partir de los años 70 (Weiner, 1990). La novedad más importante es la de conceptualizar la motivación en términos cognitivos (Corno y Snow, 1986). En la actualidad se cuenta con varias teorías que explican la motivación para el aprendizaje como, la atribucional de Weiner, (1980 y 1986); la de la autoeficacia de Bandura (1977, 1982, 1986); la de alumnos orígenes y peones (DeCharms, 1976); motivación intrínseca (Deci, 1975, Deci y Ryan, 1985), por citar algunas. A pesar de los esfuerzos realizados por distintos investigadores para integrar aspectos relacionados con la motivación académica como el Locus de control, la ansiedad o el autoconcepto, en teorías más unificadas, es necesario avanzar más en la clarificación de los constructos motivacionales, ya que la distinción entre algunos de ellos no está demasiado clara. Otro punto que necesita ser investigado es la distinción entre la motivación como rasgo o como estado, ya que algunos constructos como la orientación hacia el yo o hacia la tarea, parecen funcionar como rasgo y como estado (Nicholls, 1984, 1989; Nolen, 1988).

La mayor parte de las investigaciones con las que se cuenta son estudios corre-

lacionales o de regresión múltiple, bien para validar los constructos (Gottfried, 1990), o bien para establecer la asociación y/o el peso de distintas variables consideradas en algunos de los constructos motivacionales estudiados con el rendimiento académico o en otro tipo de tareas. En este sentido, Multon, Brown y Lent (1991) a partir de un *metaanálisis* de 39 investigaciones concluyen que hay una relación positiva entre la autoeficacia percibida, el rendimiento y la persistencia en las tareas.

Algunos trabajos de investigación empiezan a emplear constructos motivacionales diferenciales (motivación por la tarea/motivación por el yo; motivación intrínseca o extrínseca) como variables en investigaciones sobre estructuración de situaciones de aprendizaje, o sobre profundidad de procesamiento de la información, el uso de refuerzos o la utilización de estrategias cognitivas.

Pardo Merino y Alonso Tapia, (1990) estudian las diferencias entre los sujetos motivados por el yo (autovaloración) o por la tarea, en situaciones de aprendizaje competitivo y cooperativo.

Graham y Golan (1991) en dos estudios con niños de 5º y 6º grado de la escuela elemental, con un *diseño experimental*, analizan la relación entre orientación hacia el yo o hacia la tarea y el nivel de profundidad de procesamiento de la información en tareas de codificación y recuerdo de información verbal. Los resultados obtenidos muestran que los tipos de orientación motivacional tienen un efecto diferencial en la profundidad de procesamiento de la información. Nolen (1988) encuentra resultados similares con alumnos de universidad en cuanto al tipo de estrategias de estudio empleadas (superficiales por los alumnos con orientación motivacional hacia el yo, y profundas, los orientados hacia la tarea). Se requieren más investigaciones sobre el influjo de las metas de aprendizaje de los alumnos en el modo en que afrontan las tareas, del mismo modo es preciso probar la interacción entre los factores motivacionales y cognitivos, identificados por Garner (1990) y por Alonso Tapia (1991 y 1992), que han demostrado tener un peso importante en el desarrollo y uso de estrategias de aprendizaje y estudio por parte de los alumnos.

El estudio *ex post facto* de Miller y Hom (1990) analiza el uso de refuerzos y el papel que desempeñan según el tipo de orientación motivacional del sujeto hacia el yo o hacia la tarea. Los resultados de su trabajo con estudiantes de college, muestran que las recompensas externas reducen la motivación en sujetos para los que la tarea era de poco valor para el yo, y la promueven en aquellos que ven la tarea como una medida de su habilidad. Se concluye que las recompensas externas reducen la orientación a la tarea.

Una línea prometedora es la iniciada por Graham y Harris en 1989, con chicos con dificultades de aprendizaje en la que se estudia, a partir de un diseño cuasiexperimental, los efectos de distintos métodos de enseñanza (enseñanza directa, enseñanza de estrategias y enseñanza de estrategias con autorregulación) en las habilidades de composición escrita y en la autoeficacia percibida en sujetos con dificultades de aprendizaje. En los primeros trabajos realizados encuentran una relación entre las estrategias de entrenamiento con autorregulación y la autoeficacia percibida de los

alumnos. Sin embargo en un trabajo posterior (Sawyer, Graham y Harris 1992) no aparece esta relación. En ambas investigaciones se observa que hay una tendencia en los niños con dificultades de aprendizaje a puntuar alto en autoeficacia, aún antes de iniciar el entrenamiento. Estos autores concluyen que es posible que estos alumnos tiendan a sobreestimar sus posibilidades como indican los trabajos de Álvarez y Adelman (1986). Se requieren más investigaciones sobre el papel de la autoeficacia en la autoregulación del aprendizaje y el logro, en niños con dificultades de aprendizaje.

En la investigación realizada para estudiar la relación entre las dimensiones de estructuración de la clase y las estrategias instructivas del profesor, Alonso Tapia (1991) obtiene resultados similares a los que aparecen en las revisiones de Ames (1984) y Blumenfeld (1992). Este autor encuentra que el diseño de las tareas, la forma de ejercer la autoridad y el modo de evaluar caracterizan situaciones de clase diferentes que dan lugar a un tipo de estrategias instructivas-motivacionales diferenciadas por parte de los profesores según definan esa estructuración del aula.

En la actualidad hay una corriente de investigación, siguiendo la *metodología observacional* en el aula, interesada en comprobar los efectos de las estrategias empleadas por el profesor en la orientación motivacional intrínseca o extrínseca del alumno (Brophy, 1988; Brophy y otros 1983; Marshall, 1987). En estos trabajos se observa durante un tiempo a los profesores y alumnos en el aula, y se anotan las estrategias utilizadas por los profesores y el comportamiento de los alumnos, para posteriormente ver las diferencias en la conducta de los sujetos en función de las estrategias utilizadas por el profesor. En esa misma línea, Newby (1991) estudia la cantidad y el tipo de estrategias motivacionales empleadas por el profesor, así como su efecto en el grado de comportamiento centrado en la tarea manifestado por los estudiantes de los primeros cursos de la escuela elemental. Los datos se recogieron por el procedimiento de *observación no participante* y las estrategias motivacionales empleadas por los 30 profesores observados durante 4 meses se registraron y categorizaron siguiendo el modelo propuesto por Keller, (1987): atención, relevancia, confianza y satisfacción. De los resultados de esta investigación se desprende que los profesores emplean con mucha mayor frecuencia estrategias motivacionales externas de recompensa o castigo, que otras como el centrar la atención o mostrar la relevancia. También se observa que existe una correlación positiva entre el uso de estrategias de la categoría denominada por Keller como relevancia y el comportamiento centrado en la tarea por parte de los estudiantes; y una correlación negativa entre el uso de estrategias externas de recompensa o castigo y dicho comportamiento.

Los factores motivacionales y atribucionales se muestran como elementos importantes para mantener y consolidar lo aprendido. En un estudio efectuado por Borkowsky y otros (1990) los estudiantes que fueron entrenados en estrategias de aprendizaje y atribuían a esas estrategias sus logros, no sólo obtenían mejores resultados en su ejecución, sino que mantenían y generalizaban con mayor facilidad sus aprendizajes. Los estudiantes únicamente entrenados en estrategias de aprendizaje aumentaban su conocimiento declarativo sobre estas estrategias pero no su utilización.

REFERENCIAS BIBLIOGRÁFICAS

Introducción

- A.A.V.V. (1992): *Educación multicultural e intercultural*. Actas del Congreso de Educación Multicultural, celebrado en Ceuta en 1991. Granada: Impredisur.
- BARTOLOMÉ, M. (1983): *Pedagogía Diferencial. Aproximación a una ciencia*. Barcelona: Universidad Barcelona. Paper.
- CONGRESO NACIONAL DE PEDAGOGÍA (1992): *Educación Multicultural en la perspectiva de la Europa Unida*. Tomos I y II. Salamanca: Diputación Provincial.
- DE LA ORDEN, A. (1988): Conceptualización de la pedagogía diferencial, *Bordón*, 40, 4, 543-553.
- DENDALUCE, I. (1988): Contexto y criterios: hacia una tipologización de las metodologías de la investigación en Pedagogía Diferencial, *Bordón*, 40, 4, 599-611.
- JIMÉNEZ, C. (1990): Cuestiones sobre bases diferenciales de la educación. Madrid: UNED.
- JIMÉNEZ, C. (1991) (Coord.): *Lecturas de Pedagogía Diferencial*. Madrid: Dykinson.
- LÓPEZ, E. (1988): Problemas de investigación en algunos campos de Pedagogía Diferencial, *Bordón*, 40, 4, 611-631.
- MARÍN, R. (1992): Modelos de educación multicultural. En A.A.V.V. *Educación multicultural e intercultural*. (pp. 42-57). Granada: Impredisur.
- MARSHALL, G. A. (1972): Clasificaciones raciales populares y científicas. En AAVV. *Ciencia y concepto de raza*. Barcelona: Fontanella.
- SELBY, D. (1992): Educación para una sociedad multicultural: implicaciones curriculares y metodológicas. En *Actas X Congreso Nacional de Pedagogía*, (pp. 351-379). Salamanca: Diputación Provincial.
- SIGUÁN, M. (1992): *La escuela y la inmigración en la Europa de los 90*. Barcelona: Horsori.
- TEJEDOR, J. (1988): La pedagogía diferencial en su conceptualización metodológica, *Bordón*, 40, 4, 589-599.

Políticas Públicas

- AMERICAN ASSOCIATION OF COLLEGES FOR TEACHER EDUCATION (1978): *State Legislation Provisions and Practices Related to Multicultural Education*. Whashington: Autor.
- ANDERSON, E. J. (1984): *Public Polcy-Making*. New York: CBS College Publ.
- BALLART, X. (1992): *¿Cómo evaluar programas y servicios públicos? Aproximación sistemática y estudio de casos*. Madrid: M.A.P.
- BARTOLOMÉ, M. (1992): Diseños y metodologías de investigación desde la perspectiva de la educación intercultural. En X Congreso Nacional de Pedagogía. *Educación Intercultural en la Perspectiva de la Europa Unida, II*, (pp. 662-48). Salamanca: Diputación Provincial de Salamanca.
- CATTERALL, S. J. (1985) (Ed.): *Economic Evaluation of Public Programs, New Direction for Program Evaluation*, 26, junio.
- CIBULKA, G. J., Reed J. R. y Wong K. K. (1992) (Eds.): *The Politics of Urban Education in the United States*. London: The Falmer Press.
- CIBULKA, G. J. (1992): Urban Education as Field of Study: Problems of Knowledge and Power. En G. J. Cibulka, J. R. Reed y K. K. Wong: *The Politics of Urban Education in the United States*, (pp. 27-43). London: The Falmer Press.
- CIDE (1992): *Las desigualdades en el acceso a la educación en España*. Madrid: Ministerio de Educación y Ciencias.

- COLEMAN, J. S. y otros (1966): *Equality of Educational Opportunity*. Washington: Government Printing Office.
- COMISSIÓ INTERDEPARTAMENTAL DE PROMOCION DE LA DONA (1989): *Plan de Actuación del Gobierno de la Generalitat de Catalunya para la Igualdad de Oportunidades para la Mujer*. Barcelona: Generalitat de Catalunya.
- CONSEJERÍA DE EDUCACIÓN, CULTURA, DEPORTE Y JUVENTUD (1992): *Evaluación del I Plan de Igualdad de la Mujer*. Principado de Asturias: Autor.
- CRUMPTON, R. (1992): Policy Analysis of State Multicultural Education Programas. En A. C. Grant (Ed.): *Research and multicultural education*, (pp. 240-249). London: The Falmer Press.
- GOLLNICK, (1992): Multicultural education: policy and practices in teacher education. En A. C. Grant (Ed.): *Research and multicultural education*. London: The Falmer Press.
- GRANT, A. C. (1992) (Ed.): *Research and multicultural education*, (pp. 219-239). London: The Falmer Press.
- ESPÍN, J. V. (1991): Los programas de educación compensatoria. ¿Una respuesta a las diferencias cosiculares desde la educación? En C. Jiménez. *Lecturas de Pedagogía Diferencial*. Madrid: Dykinson.
- FREEDMAN, P. I. (1983): A National sample of multiethnic/multicultural education in secondary schools, *Contemporary Education*, 54, 2, 130-3.
- INSTITUTO DE LA MUJER (1990): *Plan para la Igualdad de Oportunidades de las mujeres 1988-1990*. Madrid: Ministerio de Asuntos Sociales.
- INSTITUTO DE LA MUJER (1991): *Evaluación del plan para la Igualdad de Oportunidades de las Mujeres (1988-1990)*: Madrid: Ministerio de Asuntos Sociales.
- LEONARD, D. H. (1990): The role of Policy Analysis and evaluation. En B. Perret (comp.). *Outils, Pratiques, Institutions pour Evaluer les Politiques Publiques*. (pp. 179-205). Actes du Séminaire PLAN-ENA. Avrilt-Juillet. Paris: La Documentación Française.
- LYNCH, J. (1986): *Multicultural Education: Principles and Practice*. London: Routledge and Kegand Paul.
- MCCARTHY, C. (1990): Multicultural approach to racial inequality in the United States, *Curriculum and Teaching*, 5,1-3, 25-35.
- MAYNARD, A. R. y MURNANE, J. R. (1981): The effects of negative income tax on school performance. *Evaluation Studies Review Annual*, 6, 179-92.
- MITCHELL, B. (1985): Multicultural education. Avicable component of American Education?, *Educational Research*, 937-11.
- MITCHELL B. (1988): *A National Survey of Multicultural Education*. Cheney: Western States Consulting and Evaluation Services.
- MITCHELL, D. E. (1984): Educational Policy Analysis: The state of de art, *Educational Administration Quartely*, 20, 3, 120-60.
- MULLIN, S. P. y SUMMERS, A. A. (1983): Is More Better? The Effectiveness of Spending on Compensatory Education, *Phi Delta Kappan*, 64, 5, 339-347.
- NAGEL, S. S. (1990): Bridging theory and practice in policy/program evaluation, *Evaluation and Program Planing*, 13, 275-283.
- NETHERLANDS CENTRAL BUREAU OF STATISTICS (1988): *Educational Indicators for the Basisvorming*. Voorburg: Autor.
- OECD (1992): *Une education et une formation de qualité pour tous*. Paris: Autor.
- OECD (1973): *Indicators of performance of educational systems*. Paris: Autor.
- OECD (1992): *The OECD international education indicators. A framework for analysis*. Paris: Autor.
- OLKINUORA, E. (1973): On the problem of developing educational indicators, *Acta Sociologica*, 16, 4, 284-302.

- PERRET, B. (1991) (Comp.): Outils, Pratiques, Institutions pour Evaluer les Politiques Publiques. *Actes du Séminaire PLAN-ENA. Avril-Juillet*. Paris: La Documentación Française.
- ROS, J. D. (1992): The role of Investigation in Program and Policy Development, *New Directions for Program Evaluation*, 56, 53-61.
- RIST, R. C. (1990) (Ed.): *Policy and Program Evaluation: Perspectives on design and utilization*. Brussels: International Institute of Administrative Sciences.
- SMITH, L. N. y SMITH, K. A. (1985): State-Level Evaluation Uses of Cost Analysis: a National descriptive Survey. CATTERALL, S. J. (ed.): *Economic Evaluation of Public Programs, New Direction for Program Evaluation*, 26, 83-97.
- SARIS, E. W. y BATISTA, F. (1992): *Social indicators with respect to de role of women in society*. Barcelona: ESADE.
- SLAVENBURG, J. H. (1991): Evaluation of Compensatory Education Programmes in Rotterdam, The Netherlands, *International Journal of Educational Research*, 15,1, 21-41.
- SUBBARAO, K. y RANEY, L. (1993): *Social gains from females Education. A cross National Study*. Whashington: World Bank.
- SUBIRACH, J. (1989): *Análisis de políticas públicas y eficacia de la administración*. Madrid: M.A.P.
- TROCHIN, K. W. (1983): Methodologically Based Discrepancies in Compensatory Education Evaluation, *Evaluation Studies. Review Annual*, 8, 633-670.
- UNESCO (1976): *Towards a System of educational Indicators, Division of educational policy and planning*. Paris: Autor.
- UTARBIGON, R. W. y HANUSHEK, F. A. (1992): *Educational Performance of the poor, lesson from Rural Northeast Brazil*. Oxford: University Press.
- VEGA FELIPE (1992): Principales indicadores para un programa de educación intercultural en localidades con alta presencia de residentes extranjeros. En X Congreso Nacional de Pedagogía. *Educación Intercultural en la Perspectiva de la Europa Unida, II*, (pp. 720-21). Salamanca: Diputación provincial de Salamanca.
- VIGHI, I. (1992): Programas europeos de acción en favor de las mujeres. *Salud y Trabajo*, 79, 4-7.
- WOODHALL, M. (1992): *Cost-benefit Analysis in educational planning*. Paris: Unesco.

Diferencias étnicas y socioculturales: modelos cuantitativos

- BAKER, D. P. (1993): Compared to Japan, the U.S. is a low achiever... really. New evidence and comment on Westbury, *Educational Researcher*, 22, 3, 18-20.
- BARTOLOMÉ, M. (1992): Diseños y metodologías de investigación desde la perspectiva de la educación multicultural. En X Congreso Nacional de Pedagogía. *Educación intercultural en la Perspectiva de la Europa Unida, II*, (pp. 647-674). Salamanca: Diputación provincial de Salamanca.
- BEAUTRY, J. S. (1992): Synthesizing research in multicultural teacher education: finding and issues for evaluation of cultural diversity. En A. M. Madiso (Ed.): *Minority Issues in Program Evaluation, New directions for Program Evaluation*, 53, 69-86. San Francisco: Jossey-Bass publishers.
- CALVO BUEZAS, T. (1987): *Tratamiento de las minorías étnicas en los libros de textos escolares*. Intervención en las VII Jornadas de Enseñantes con Gitanos. Badalona.
- CALVO BUEZAS, T. (1989): *Los racistas son los otros*. Madrid: Popular.
- CALVO BUEZAS, T. (1990): *El racismo que viene*. Madrid: Tecnos.
- CARTER, R. T. (1991): Racial identity attitudes and psychological functioning, *Journal of Multicultural Counseling and Development*, 19, 105-114.
- CAZDEN, C. B. (1990): Differential treatment in New Zealand: reflections on research in minority education, *Teaching and Teacher Education*, 6, 4, 291-303.

- CIRES (1993): *La realidad social en España 1991-92*. Barcelona: Ediciones B.S.A. Coedición de Fundación BBV.
- CLIFTON, R. A., WILLIAMS, T. y CLANCY, J. (1991): The academic attainment of ethnic groups in Australia: a social psychological model, *Sociology of Education*, 64, 111-126.
- COLECTIVO IOÉ, (1992): La inmigración extranjera en España: Sus características diferenciales en el contexto europeo. *III Jornadas Estatales Sobre Perspectivas del Trabajo Social: Inmigración-Emigración*, (paper).
- DAVIS, J. E. (1992): Reconsidering the use of race as an explanatory variable in program evaluation. En A. M. Madison (Ed.): «Minority Issues in Program Evaluation». New directions for Program Evaluation. N° 53. (pp. 55-67). San Francisco: Jossey-Bass.
- DUNCAN, C. y PRYZWANSKY, W. B. (1993): Effects of race, racial identity development, and orientation style on perceived consultant effectiveness, *Journal of Multicultural Counseling and Development*, 21, 2, 88-96.
- DUNN, CH. W. y TUCKER, C. M. (1993): Black children's Adaptive Functioning and Maladaptive behavior associated with quality of family support, *Journal of Multicultural Counseling and Development*, 21, 2, 79-87.
- ENGSTROM, C. M. y SEDLACEK, W. E. (1991): A study of prejudice toward university student-athletes, *Journal of Counseling and Development*, 70, 189-193.
- FREDRICKSON, R. H., LIN, J. G., y XING, S. (1992): Social status ranking of occupations in the people's Republic of China, Taiwan, and the United States, *The Career Development Quarterly*, 40, 351-360.
- GIM, R. H., ATKINSON, D. R., y KIM, S. J. (1991): Asian-American Acculturation, counselor ethnicity and cultural sensitivity, and rating of counselors, *Journal of Counseling Psychology*, 38, 1, 57-62.
- GRAND, C. A. y MILLAR, S. (1992): Research and Multicultural Education: barriers, need and boundaries. En C. A. Grant (Ed.): *Research and multicultural education*. London: The Falmer Press.
- HELMS, J. E. (1984): Toward a theoretical explanation of the effects of race on counseling: A black and white model, *The Counseling Psychologist*, 12, 4, 153-156.
- HELMS, J. E., y CARTER, R. T. (1991): Relationships of white and black racial identity attitudes and demographic similarity to counselor preferences, *Journal of Counseling Psychology*, 38, 4, 446-457.
- HENRY, P., BARDO, H. R., Y HENRY, C. A. (1992): The effectiveness of career development seminars on African American premedical students: a program evaluation using the Medical Career Development Inventory, *Journal of Multicultural Counseling and Development*, 20, 99-112.
- HESS, R. S., y STREET, E. M. (1991): The effect of acculturation on the relationship of counselor ethnicity and client ratings, *Journal of Counseling Psychology*, 38, 1, 71-75.
- HILLIARD, A. G. (1989): Kemet (Egyptian) Historical Revision: Implications for Cross-Cultural Evaluation and Research in Education, *Evaluation Practice*, 10, 7-21.
- HLEBOWITSH, P. S., y TÉLLEZ, K. (1993): Pre-service Teachers and their students: early views of race, gender and class, *Journal of Education for Teaching*, 19, 1, 41-52.
- HOUSE, E. R. (1990): Methodology and Justice. En K. A. Sirotnik (Ed.): *Evaluation and social justice: Issues in Public Evaluation. New Directions for Program Evaluation*, 45. (pp. 23-36): San Francisco: Jossey Bass.
- KYSEL, F. (1988): Ethnic Background and examination results, *Educational Research*, 30, 2, 83-89.
- LUCAS, M. S. (1993): Personal, social, academic, y career problems expressed by minority college students, *Journal of Multicultural Counseling and Development*, 21, 1, 2-13.
- LADSON-BILLIG, G. (1992): Culturally relevant teaching: the key to making multicultural education work. En C. A. Grant (Ed.): *Research and multicultural education*. (pp. 106-121). London: The Falmer Press.
- LINCOLN, Y. S. (1991): The arts and sciences of program evaluation, *Evaluation Practice*, 12, 1-7.

- MADISON, A. M. (1992): Primary inclusion of culturally diverse minority program participans in the evaluation process. En A. M. Madison (Ed.): *Minority Issues in Program Evaluation. New directions for Program Evaluation*, 53. (pp. 35-43). S. Francisco: Jossey-Bass.
- MCDOWELL, C. L. (1992): Standardized tests and Program Evaluation: inappropriate measures in critical times». En A. M. Madison (Ed.): *Minority Issues in Program Evaluation. New directions for Program Evaluation*, 53. (pp. 45-54). San Francisco: Jossey-Bass.
- MIO, J. S. y IWAMASA, G. (1993): To do, or no to do: that is the question for white cross-cultural researchers, *The Counseling Psychologist*, 21, 2, 197-212.
- MORRIS, J. R. (1991): Racial attitudes of undergraduates in greek housing, *The College Student Journal*, 25, 1, 501-505.
- PARHAM, T. A., y HELMS, J. E. (1985a): Relation of racial identity attitudes to self-actualization and affective states of black student, *Journal of Counseling Psychology*, 32, 2, 431-440.
- PARHAM, T. A., y HELMS, J. E. (1985b): Attitudes of racial identity and self-esteem: An exploratory investigation, *Journal of College Student Personnel*, 26, 143-146.
- PARKER, W. M. y LORD, S. L. (1993): Characteristics of race Role Models for young African American men: An exploratory survey, *Journal of Multicultural Counseling and Development*, 21, 2, 97-105.
- PATTON, M. Q. (1987): *How to use Qualitative Methods in Evaluation*. London: Sage.
- PEDERSEN, P. (1987): Ten frequent assumption of cultural bias in Counseling, *Journal of Multicultural Counseling and Development*, 15, 1, 16-24.
- PHELPS, R. E., MEARA, N. M., DAVIS, K. L., y PATTON, M. J. (1991): Blacks' and whites' perceptios of verbal aggression, *Journal of Counseling and Development*, 69, 345-350.
- PONTEROTTO, J. G. (1991): The nature of prejudice revisited: implications for counseling interventions, *Journal of Counseling and Development*, 70, 216-224.
- POWER, K. M., DOLLY, J. P., y BLAINE, D. D. (1993): The effects of overgeneralizing ethnic categories in educational evaluation, *Studies in Educational Evaluation*, 19, 15-23.
- PRADA, M. A. DE, PEREDA, C. y ACTIS, W. (1992): Colectivo Ioé. Extranjeros y españoles. Mas allá de opiniones y actitudes. Aproximación a la lógica discursiva de las relaciones. *Racismo y educación: Hacia una educación multicultural*. Universidad de verano Antonio Machado. Baeza.
- PRICE, G. G. (1992): Using quantitative methods to explore Multicultural Education. En C. A. Grant (Ed.): *Research and multucultural education*. London: The Falmer Press.
- PRIEST, R. (1991): Racism and prejudice as negative impacts on african american clients therapy, *Journal of Counseling and Development*, 70, 213-215.
- RAMSEY, P. G. (1991): The salience of race in young children growing up in an all-white community, *Journal of Educational Psychology*, 83, 1, 28-34.
- RUMBERGER, R. W. y WILLMS, J. D. (1992): The impact of racial and ethnic segregation on the achievement gap in California high schools, *Educational Evaluation and Policy Analysis*, 14, 4, 377-396.
- SÁNCHEZ, I., ROMERA, M. J. y SÁEZ, R. (1992): Información bibliográfica y documental sobre educación multicultural y educación intercultural. En *X Congreso Nacional de Pedagogía. Educación intercultural en la Perspectiva de la Europa Unida*. II. (pp. 805-841). Salamanca: Diputación provincial de Salamanca.
- SERGEANT, M. T., WOODS, P. A. y SEDLACEK, W. E. (1992): University student attitudes toward arabs: intervention implications, *Journal of Multicultural Counseling and Development*, 20, 123-131.
- SLEETER, C. y GRANT, C. (1987): An analysis of multicultural education in the United States», *Harvard Educational Review*, 57, 412-444.
- SODOWSKY, G. R., WAI MING, E., y PLAKE, B. (1991): Moderating effects of sociocultural variables on Acculturation Attitudes of Hispanics and Asian Americans, *Journal of Counseling and Development*, 70, 1, 194-205.

- STOCKDILL, S. H., DUHON-SELLS, R. M., OLSON, R. A. y PATTON, M. Q. (1992): Voices in the design and evaluation of a multicultural education program: a developmental approach. En A. M. Madison (Ed.): *Minority Issues in Program Evaluation. New directions for Program Evaluation*, 53. (pp. 17-33). San Francisco: Jossey-Bass publishers.
- SUE, D. W. (1981): *Counseling the culturally different: theory and practice*. New York: Wiley.
- SUTTER, J. A. y MCCAUL, E. J. (1993): Issues in cross-cultural counseling: An examination of the meaning and dimensions of tolerance, *International Journal for the Advancement of Counseling*, 16, 3-18.
- TAMIR, P. (1992): Ethnic origin, socioeconomic status and science learning outcomes of high school students in Israel, *Curriculum and Teaching*, 7, 1, 27-35.
- THIRKELL, B. y WORRALL, N. (1989): Differential ethnic bias in Bengali and white children, *Educational Research*, 31, 3, 181-187.
- TOKAR, D. M., y SWANSON, J. L. (1991): An investigation of the validity of Helms's (1984) Model of white racial identity development, *Journal of Counseling Psychology*, 38, 3, 296-301.
- SIRAJ-BLATCHFORD, I. (1991): A study of black students' perceptions of racism in Initial Teacher Education, *British Educational Research Journal*, 17, 1, 35-50.
- WADE, P. y BERSTEIN, B. L. (1991): Culture sensitivity training and counselor's race: effects on black female clients' perceptions and attrition, *Journal of Counseling Psychology*, 38, 1, 9-15.
- WESTBURY, I. (1992): Comparing American and Japanese achievement: Is the United States really a low achiever?, *Educational Researcher*, 21, 5, 18-24.
- WESTBURY, I. (1993): American and Japanese achievement... again. A response to Baker, *Educational Researcher*, 22, 3, 21-26.
- YOUNG, J. W. (1991): Improving the prediction of performance of ethnic minorities using the IRT-based GPA, *Applied Measurement in Education*, 4, 3, 229-239.
- ZAPF, M. K. (1991): Cross-cultural transitions and wellness: dealing with culture shock, *International Journal for the Advancement of Counseling*, 14, 105-119.

Diferencias étnicas y socioculturales: modelos cualitativos

- AGUADO, M. T. (1991): «La educación intercultural: Concepto, paradigmas, realizaciones. En C. Jiménez (Coord.): *Lecturas de Pedagogía Diferencial*. (pp. 87-105). Madrid: Dykinson.
- ALEGRE, O. M. (1990): «El entrenamiento musical y la lateralidad en niños deficientes visuales», *Curriculum*, 1, 131-142.
- ANDERSON, G. L. (1989): «Critical Ethnography in Education: Origins, Current Status and New Directions», *Review of Educational Research*, 59, 3, 249-270.
- ASANTE, M. K. (1991): «Multiculturalism: An Exchange», *The American Scholar*, 60, 2, 267-272.
- BARTOLOMÉ, M. y ANGUERA, M. T. (1990): *La investigación cooperativa: vía para la innovación en la Universidad*. Barcelona: PPU.
- BARTOLOMÉ, M. (1991): Diferencias socioculturales y marginación: aproximación educativa al problema. En C. Jiménez (Coord.): *Lecturas de Pedagogía Diferencial*. (pp. 53-87). Madrid: Dykinson.
- BARTOLOMÉ, M. y ACOSTA, A. (1991): Articulación de la educación popular y la educación formal. Ponencia presentada en las *Jornadas de Investigación Participativa*, celebradas en la Universidad de Barcelona, 15-20 de Sep. (paper).
- BARTOLOMÉ, M. (1992): Diseños y metodologías de investigación desde la perspectiva de la educación multicultural. Ponencia presentada en el *X Congreso nacional de Pedagogía sobre Educación Intercultural en la Perspectiva de la Europa Unida*. II. (pp. 647-674). Salamanca: Diputación Provincial de Salamanca.
- BARTOLOMÉ, M., CABRERA, F., ESPÍN, J. V., MARÍN, M. A., MARTÍN, J., DEL RINCÓN, D. y

- RODRÍGUEZ, M. (1992): *Diagnóstico de las diferencias étnicas y de los procesos desarrollados en la educación Primaria*. Proyecto de investigación subvencionado por el CIDE. Barcelona: Universidad de Barcelona.
- BORMAN, K. M., TIMM, P., EL-AMIN, Z. y WINSTON, M. (1992): Using Multiple Strategies to Assess Multicultural Education in a School District. En C. A. Grant (1992): *Research & Multicultural Education*, (pp. 71-88): London: The Falmer Press.
- BROWN, S. C. (1992): Application of Multicultural and Global Concepts in Preservice Teachers' Classrooms: A Preliminary Report, *Curriculum*, 13, 2, 99-102.
- CARRINGTON, B. (1983): Sport as a side-track: an analysis of West Indian involvement in extracurricular sport. En L. Barton y S. Walker (Eds.): *Race, Class and Education*, Beckenham: Croom Helm.
- CARRINGTON, B. y WOOD, E. (1983): Body talk, *Multiracial Education*, 11, 2, 29-38.
- CLAUX, R. y LEMAY, P. (1991): La Recherche-Action: Fondaments, pratique et formatio. L'actualité de Kurt Lewin. Ponencia presentada en el *Colloque: La Recherche-Action de Kurt Lewin aux Pratiques Québécoises contemporaines*. 59 Congress de l'ACFAS. Université de Sherbrooke. 21-24 Mai 1991. (paper).
- CODINA, M. T. (1991): Educación en la diversidad, *Cuadernos de Pedagogía*, 196, 27-28.
- DENIA, N., GIRÓ, J. y NAVARRO, J. M. (1988): *Elements per a una reflexió didàctica sobre l'etnocentrisme als manuals escolars d'història*. Barcelona: PPU.
- DUNKIN, M. J. y DOENAU, S. J. (1989): Etnia y comportamiento en el aula. En T. Husen y T. N. Postlethwaite (Eds.): *Enciclopedia Internacional de la Educación*, (pp. 2.487-2.490). Barcelona: M.E.C.: Vicens Vives.
- ESPÍN, J. V. (1991): Los programas de educación compensatoria. ¿Una respuesta a las diferencias socioculturales desde la educación? En C. Jiménez (Coord.): *Lecturas de Pedagogía Diferencial*, (pp. 105-135). Madrid: Dykinson.
- ETXEARRIA, F. (1992): Educación intercultural, racismo y europeísmo. En Varios *Educación multicultural e intercultural*. Actas del Congreso de Educación Multicultural celebrado en Ceuta en 1991. Granada. Impredisur.
- EVERTSON, C. M. y GREEN, J. (1989): La observación como indagación y método. En M. C. Wittrock: *La investigación de la enseñanza, II. Métodos cualitativos y de observación*. Barcelona: Paidós: MEC.
- GARCÍA DE LEÓN, M. A. (1990): Nota de investigación: sociologías ricas/sociologías pobres: el caso del etnocentrismo de la sociología norteamericana, *Tempora*, 15-16, 75-80.
- GHOSH, R. y TARROW, N. (1993): Multiculturalism and Teacher Educators: views from Canada and USA, *Comparative Education*, 29, 1, 81-92.
- GIBBS, J. T. (1990): Biracial adolescents. En J. T. Gibbs y L. N. Huang (Eds.): *Children of color*, (pp. 332-350). San Francisco: Jossey-Bass.
- GORDON, B. M. (1992): The Marginalized Discourse of Minority Intellectual Thought in Traditional Writings on Teaching. En C. A. Grant (Ed.): *Research & Multicultural Education*, (pp. 19-31): London: The Falmer Press.
- GRANT, C. A. (1992): *Research & Multicultural Education*. London: The Falmer Press.
- GRANT, C. A. y MILLAR, S. (1992): Research and Multicultural Education: Barriers, Needs and Boundaries. En C. A. Grant (1992): *Research & Multicultural Education*, (pp. 7-18). London: The Falmer Press.
- GRUGEON, E. y WOODS, P. (1990): *Educating all. Multicultural perspectives in the Primary School*. London: Routledge.
- HARRY, B. (1992): An Ethnographic Study of Cross-Cultural Communication With Puerto Rican-American Families in the Special Education System, *American Educational Research Journal*, 29, 3, 471-494.
- HERRING, R. D. (1992): Biracial Children: An Increasing Concern for Elementary and Middle School Counselor, *Elementary School Guidance and Counseling*, 27, 2, 123-130.

- JORGENSEN, D. (1989): *Participant Observation: A Methodology for Human Studies*. California: Sage.
- KEEVES, J. P. (1988): Sex Differences in Ability and Achievement. En J. P. Keeves (Ed.): *Educational Research, Methodology, and Measurement: An International Handbook*, (pp. 482-487). Oxford: Pergamon Press.
- KNAPP, M. S. (1979): Contribuciones etnográficas a la investigación evaluativa. En T. D. Cook y CH. S. Reichardt: *Qualitative and Quantitative Methods in Evaluation Research*. London: Sage. (Trad. cast.: *Métodos cualitativos y cuantitativos en investigación evaluativa*. Madrid: Morata).
- LEWIS, A. (1992): Group Child Interviews as a Research Tool, *British Educational Research Journal*, 18, 4, 413-421.
- LINCH, J. (1992): *Education for Citizenship in a Multicultural Society*. New York: Cassell.
- LUCAS, M. S. (1993): Personal, Social, Academic, and Career Problems Expressed by Minority College Students, *Journal of Multicultural Counseling and Development*, 21, 2-13.
- MARSHALL, C. y ROSSMAN, G. (1989): *Designing qualitative research*. Londres: Sage.
- MATEOS, F. C. (1992): La investigación-acción para una educación multicultural. Propuesta de intervención en una comunidad gitana. En Varios. *Educación multicultural e intercultural*. Actas del Congreso de Educación Multicultural celebrado en Ceuta en 1991. Granada: Impredisur.
- MAXWELL, J. A., BASHOOK, P. G. y SANDLOW, L. J. (1986): Combining Ethnographic and Experimental Methods in Educational Evaluation. A Case Study. En D. Fetterman: *Ethnography, step by step*, (pp. 121-143). Beverly Hills, Ca: Sage.
- MAW, J. (1991): Ethnocentrism, history textbooks and teaching strategies: presenting the URSS. *Research Papers in Education*, 6, 3, 153-169.
- MIRA, E. y ROIG, M. (1991): El aprendizaje de la comunicación. *Cuadernos de Pedagogía*, 196, 20-26.
- MONTERO-SIEBURTH, M. y GRAY, C. A. (1992): Riding the Wave: Collaborative Inquiry Linking Teachers at the University and the Urban High School. En C. A. Grant (1992): *Research & Multicultural Education*, (pp. 122-140). London: The Falmer Press.
- NILES, F. S. (1993): Issues in Multicultural Counselor Education, *Journal of Multicultural Counseling and Development*, 21, 14-21.
- NORQUAY, N. (1990): Life History Research: memory, schooling and social difference, *Cambridge Journal of Education*, 20, 3, 291-300.
- OGBU, J. U. (1992): Understanding Cultural Diversity and Learning, *Educational Researcher*, 21, 8, 5-14.
- PATE, R. H. y BONDI, A. M. (1992): Religious Beliefs and Practice: An Integral Aspect of Multicultural Awareness, *Counselor Education and Supervision*, 32, 108-115.
- PRICE, G. G. (1992): Using Quantitative Methods to Explore Multicultural Education. En C. A. Grant (1992): *Research & Multicultural Education*, (pp. 58-70). London: The Falmer Press.
- QUICKE, J. (1991): Social background, Identity and Emergent Political Consciousness in the Sixth Form, *Cambridge Journal of Education*, 21, 1.
- RAMISETTY-MIKLER, S. (1993): Asian Indian Immigrants in America and Sociocultural Issues in Counseling, *Journal of Multicultural Counseling and Development*, 21, 36-49.
- RAMLER, S. (1991): Global Education for 21st Century, *Educational Leadership*, 48, 7, 44-46.
- RAVITCH, D. (1991): Multiculturalism: An Exchange, *The American Scholar*, 60, 2, 272-276.
- ROBINSON, T. (1992): Transforming at-risk educational practices by understanding and appreciating differences, *Elementary School Guidance and Counseling*, 27, 2, 84-95.
- SLEETER, C. L. y GRANT, C. A. (1987): An analysis of multicultural education in the United States, *Harvard Educational Review*, 57, 4, 421-44.
- SMOCK, P. J. y WILSON, F. D. (1991): Desegregation and the Stability of White Enrollments: A School-level Analysis, 1968-84, *Sociology of Education*, 64, 278-292.
- SOBOL, T. (1990): Understanding Diversity, *Educational Leadership*, 48, 3, 27-30.

- SPRADLEY, J. P. (1979): *The Ethnography Interview*. Nueva York: Holt, Rinehart and Winston.
- SPRADLEY, J. P. (1980): *Participant Observation*. Nueva York: Holt, Rinehart and Winston.
- STANFIELD II, J. H. (1993a): Methodological Reflections. An Introduction. En J. H. Stanfield II y R. M. Dennis (Eds.): *Race and Ethnicity in Research Methods*, (pp. 3-15). London: Sage.
- STANFIELD II, J. H. (1993b): Epistemological Considerations. En J. H. Stanfield II y R. M. Dennis (Eds.): *Race and Ethnicity in Research Methods*, (pp. 16-36). London: Sage.
- STANFIELD II, J. H. y DENNIS, R. M. (Eds.) (1993): *Race and Ethnicity in Research Methods*. London: Sage.
- SUTTLES, G. D. (1986): School Desegregation and 'National Community. En J. Prager, D. Logshore y M. Seeman (Ed.): *School Desegregation Research*. New York: Plenum Press.
- TAYLOR, S. y BOGDAN, R. (1986): *Introducción a los métodos cualitativos de investigación*. Buenos Aires: Paidós.
- VAN DIJK, T. (1993): Analyzing Racism Through Discourse Analysis. Some Methodological Reflections. En J. H. Stanfield II y R. M. Dennis (Eds.): *Race and Ethnicity in Research Methods*, (pp. 92-134). London: Sage.
- VARIOS (1991): *Educación multicultural e intercultural*. Actas del Congreso de Educación Multicultural celebrado en Ceuta. Granada: Impredisur.
- WEIS, L. (1992): Reflections on the Researcher in a Multicultural Environment. En C. A. Grant. *Research & Multicultural Education*, (pp. 47-57). London: The Falmer Press.
- WOODS, P. y GRUGEON, E. (1990): Pupils and «Race»: integration and disintegration in primary schools, *British Journal of Sociology of Education*, 11, 3, 309-326.
- WRIGTH, C. (1986): School processes: an ethnographic study. En J. Eggleston, D. Dunn y M. Anjali: *Education for Some: the educational and vocational experiences of 15-18 year old members of minority ethnic groups*. Stoke-on-Trent: Trentham Books.
- ZIMMERMAN, J. E. y SODOWSKY, G. R. (1993): Influences of Acculturation on Mexican-American Drinking Practices: Implications for Counseling, *Journal of Multicultural Counseling and Development*, 21, 22-35.

Diferencias de género

- ACHER, J. y MCDONALD, M. (1991): Gender roles and school subjects in adolescent girls, *Educational Research*, 33, 1, 55-64.
- ALBERDI, I. (1987): Coeducación y sexismo en la enseñanza media. En Instituto de la Mujer (Ed.): *La investigación en España sobre mujer y educación*, (pp. 26-32). Madrid: Instituto de la Mujer.
- ALTING, A. y PELGRUM, W. J. (1990): The Siss in The Netherlands: Descriptives and Gender Differences, *Studies in Educational Evaluation*, 16, 421-441.
- ALVARO, I. y MONGE, I. (1984): La familia en los libros de texto de lectura infantil, *Revista de Educación*, 275.
- ARNESEN, A. L. (1992): *Políticas y estrategias en Noruega para la promoción de la igualdad de oportunidades en la formación del profesorado*. II Jornadas de Formación Inicial del Profesorado en Educación no sexista, febrero, (paper).
- ASKEW, S. y ROSS, C. (1991): *Los chicos no lloran. El sexismo en educación*. Barcelona: Paidós.
- BANDURA, A. (1986): *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, Nueva York: Prentice-Hall.
- BARTOLOMÉ, M. (1991): *Trayectoria educativa de la mujer: la lucha inacabada*. Ponencia presentada al Simposio sobre «La Mujer, nueva realidad, respuestas nuevas». Sevilla, 27-29 de diciembre, (paper), 34 pp.

- BELENKY, M. F. y otros (1986): *Women's ways of Knowing: The development of self, voice, and mind*. Nueva York: Basic Books.
- BENNETT, R. E., GOTTESMAN, R. L., Rock, D. A. y Cerullo, F. (1993): Influence of Behavior Perceptions and Gender on Teachers «Judgments of Students» Academic Skill, *Journal of Educational Psychology*, 85, 2, 347-356.
- BETZ, N. E. y HACKETT, G. (1981): The relationship of career-related self-efficacy expectations to perceived career options in college women and men, *Journal of Counseling Psychology*, 28, 399-410.
- BIKLEN, S. y SHAKESHAFT, CH. (1985): *Handbook for achieving sex equity through education*. Baltimore: The John Hopkins University Press.
- BIRENBAUM, M. y KRAEMER, R. (1992): «Effects of gender and ethnicity on students» perceptions of mathematics and language study, *Journal of Research and Development in Education*, 26, 1, 30-37.
- BROPHY, J. E. (1985): Interactions of Male and Female Students with Male and Female Teachers. En L. C. Wilkinson y C. B. Marrett (Eds.): *Gender influences in classroom Interaction*. Londres: Academic Press.
- BROTONS, P. (1987): Coeducación: imágenes y roles. En Instituto de Ciencias de la Educación (Ed.): *Mujer y educación. El sexismo en la enseñanza*, (pp. 145-160): Barcelona: ICE de la Universidad Autónoma.
- BURLEW, A. K. y JOHNSON, J. L. (1992): Role Conflict and Career advancement among African American Women in nontraditional professions, *The Career Development Quarterly*, 40, 302-312.
- CALZADA, T. E. (1987): Algunos aspectos de la «feminización» del trabajo docente. Una observación en la Escuela de Magisterio. En Instituto de Ciencias de la Educación (Ed.), *Mujer y educación. El sexismo en la enseñanza*, (pp. 103-107). Barcelona: ICE de la Universidad Autónoma.
- CAMPBELL, P. (1981): *The Impact of Societal Biases on Research Methods*. Washington: Paper for The National Institute of Education.
- CAREAGA, P. (1987): Modelos masculinos y femeninos en los libros de texto de E.G.B. En Instituto de la Mujer (Ed.): *La investigación en España sobre mujer y educación*, (pp. 34-40). Madrid: Instituto de la Mujer.
- CASELLES, J. F. (1991): Pedagogía diferencial, sexo y educación: El sexismo en el sistema educativo mixto y la coeducación como alternativa. En C. Jiménez Fernández (Coord.): *Lecturas de Pedagogía Diferencial*, (pp. 135-170). Madrid: Dykinson.
- CASEY, K. y APPLE, M. W. (1992): El género y las condiciones de trabajo del profesorado: el desarrollo de su explicación en America, *Educación y Sociedad*, 10, 7-21.
- CINI, A. C. y BAKER, S. B. (1987): Enhancing rural female adolescents' awareness of nontraditional careers. *The Career Development Quarterly*, 35, 316-325.
- COSPER, W. (1970): *An analysis of sex differences in teacher-student interaction as manifest in verbal and nonverbal behavior uses*. Unpublished Ed. D. dissertation, The University of Tennessee.
- CRAMER, S. H.; WISE, P. S. y COLBURN, E. D. (1977): An evaluation of a treatment to expand the career perceptions of junior high school girls, *School Counselor*, 25, 124-129.
- CRUMP, S. J. (1990): Gender and Curriculum: power and being female, *British Journal of Sociology of Education*, 11, 4, 365-385.
- DAVINSON, J. y KANYUKA, M. (1992): Girls' Participation in Basic Education in Southern Malawi, *Comparative Education Review*, 36, 4, 446-466.
- DENDALUCE, I. (1987): Líneas de investigación educativa empírico-experimental en la Universidad del País Vasco, *Revista de Investigación Educativa*, 5, 9, 49-77.
- DENDALUCE, I. (1988): Contexto y criterios hacia una tipologización de las metodologías de la investigación en Pedagogía Diferencial, *Bordón*, 40, 4, 599-610.
- DONN, M. (1986): The position of women in education, 1985. *Teacher Career and Promotion Study*, Department of Education, Wellington, Nueva Zelanda.

- DONN, M. (1987): The position of women in education, 1986. *Teacher Career and Promotion Study*, Department of Education, Wellington, Nueva Zelanda.
- EICHLER, M. (1991): *Nonsexist research methods. A practical guide*. London: Routledge, Chapman and Hall, Inc.
- ESCOLÀ, P. y otros (1991): Projecte d'intervenció en coeducació, *Guix*, 161, 11-16.
- ESPÍN, J. V. y FIGUERA, P. (1991): *Las expectativas de rol en universitarios/as*. Comunicación presentada al Simposio sobre «La Mujer, nueva realidad, respuestas nuevas». Sevilla, 27-29 de diciembre, (paper).
- FARMER, H. S. (1985): Model of Career and Achievement Motivation for Women and Men, *Journal of Counseling Psychology*, 32, 3, 363-390.
- FASSINGER, R. E. (1990): Causal models of career choice in two samples of college women, *Journal of Vocational Behavior*, 36, 225-248.
- FAULKNER, J. (1991): Mixed-sex schooling and equal opportunity for girls: a contradiction in terms? *Research Papers in Education*, 6, 3, 197-223.
- FEMINARIO DE ALICANTE (1987): *Elementos para una educación no sexista*. Valencia: Víctor Orenge.
- FRIED, L. (1989): ¿Se trata en el jardín de infancia a las niñas igual que a los niños? Análisis de conversaciones entre educadoras y niños. *Revista de Educación*, 290, 197-214.
- GALLOS, J. V. (1989): Exploring women's development: Implications for career theory, practice and research. En M. B. Arthur, D. T. Hall y B. S. Lawrence (Eds.): *Handbook of career theory*, (pp. 110-132): Cambridge: Cambridge University Press.
- GARRETA, N. (1984): La presencia de la mujer en los textos escolares, *Revista de Educación*, 275.
- GEWIRTZ, D. (1991): Analyses of Racism and Sexism in Education and Strategies for Change, *British Journal of Sociology of Education*, 12, 2, 183-201.
- GILKES, V. T. (1990): «Liberated to Work life Dogs!»: Labeling Black women and their work. En H. Y. Grossman y N. L. Chester (Eds.): *The experience and meaning of work in women's lives*, (pp. 165-188). Hillsdale, Nueva York: Erlbaum.
- GILLIGAN, C. (1979): Woman's place in man's life cycle, *Harvard Educational Review*, 49, 431-446.
- GOMBOS-TOZZO, S. y GOLUB, S. (1990): Playing Nurse and Playing Cop: Do They Change Children's Perceptions of Sex-Role Stereotypes?, *Journal of Research in Childhood Education*, 4, 2, 123-129.
- GUTTMAN, M. A. (1991): Issues in the Career Development of Adolescent Females: Implications for Educational and Guidance Practices, *Guidance and Counseling*, 6, 3, 59-74.
- HALL, J. y COOPER, J. (1991): Gender, experience and attributions to the computer, *Journal Educational Computing Research*, 7, 1, 51-60.
- HALSTEAD, M. (1991): Radical feminism, Islam and the single-sex school debate, *Gender and Education*, 3, 3, 263-278.
- HARDING, S. y HINTIKKA, M. B. (1983): *Discovering reality: Feminist perspectives on epistemology, metaphysics, methodology, and philosophy of science*. Dordrecht, Holland: Reidel.
- HAW, K. F. (1991): Interactions of gender and race-a problem for teachers? A review of the emerging literature, *Educational Research*, 33, 1, 12-21.
- HEDGES, L. V. y FRIEDMAN, L. (1993): Gender Differences in Variability in Intellectual Abilities: A Reanalysis of Feingold's Results, *Review of Educational Research*, 63, 1, 94-105.
- HELD, J. D.; ALDERTON, D. L.; FOLEY, P. P. y SEGALL, D. O. (1993): Arithmetic reasoning gender differences: Explanations found in the armed services vocational aptitude battery, *Learning and Individual Differences*, 5, 2, 171-186.
- HERAS, P. (1987): El papel de la mujer en la enseñanza y en los libros de texto en Cataluña. En Instituto de la Mujer (Ed.): *La investigación en España sobre mujer y educación*, (pp. 42-50). Madrid: Instituto de la Mujer.

- HOROWITZ, T. (1992): Gender plus-aggression, gender and other social variables, *Education and Society*, 10, 1, 63-71.
- INSTITUTO DE LA MUJER (1991): *I Jornadas sobre Formación inicial del profesorado en Educación no sexista*. Madrid: Autor.
- INSTITUTO DE LA MUJER (1992): *II Jornadas sobre Formación inicial del profesorado en Educación no sexista*. Madrid: Autor.
- JIMÉNEZ ALEIXANDRE, M. P. (1992): Una perspectiva no sexista en la formación del profesorado de ciencias. En Instituto de la Mujer. *II Jornadas sobre Formación inicial del profesorado en Educación no sexista*. Madrid: Instituto de la Mujer.
- JIMÉNEZ FERNÁNDEZ, C. (1990): *Cuestiones sobre bases diferenciales de la educación*. Madrid: UNED.
- JOHNSON, J. Y EXTEMA, J. S. (1982): *Positive images: Breaking stereotypes with children's television*. Beverly Hills, CA: Sage.
- KAISER, A. (1991): West Germany. En M. Wilson (Ed.): *Girl and Young Women in Education. A European Perspective*, (pp. 187-202). Oxford: Pergamon Press.
- KELLER, E. F. (1985): *Reflections on Gender and Science*. New Haven: Yale University Press.
- KELLY, G. P. (1989): Nuevas orientaciones en la investigación de la educación y de la mujer en el Tercer Mundo. El desarrollo de los enfoques centrados en la mujer. *Revista de Educación*, 290, 59-80.
- KELLY, J. (1991): A Study of Gender Differential Linguistic Interaction in the Adult Classroom, *Gender and Education*, 3, 2, 137-143.
- KING, S. (1989): Sex differences in a Causal Model of Career Maturity, *Journal of Counseling and Development*, 68, 2, 208-215.
- LAFRANCE, M. (1991): School for Scandal: different educational experiences for females and males, *Gender and Education*, 3, 1, 3-12.
- LEMING, J. S. (1992): The Influence of Contemporary Issues Curricula on School-Aged Youth, *Review of Research in Education*, 18, 111-161.
- LEWIS, M. (1990): Interrupting Patriarchy: Politics, Resistance, and Transformation in the Feminist Classroom, *Harvard Educational Review*, 60, 4, 467-488.
- LÓPEZ, P. y SOBRÉ, M. (1993): Talleres de orientación, *Cuadernos de Pedagogía*, 212,
- LOW, R. y OVER, R. (1993): Gender Differences in Solution of Algebraic Word Problems Containing Irrelevant Information, *Journal of Educational Psychology*, 85, 2, 331-339.
- MAHER, F. A. y RATHBONE, Ch. H. (1989): La formación del profesorado y la teoría feminista. Algunas implicaciones prácticas, *Revista de Educación*, 290, 93-112.
- MARCO, A. (1992): Estereotipos sexistas en diccionarios escolares. En Instituto de la Mujer. *II Jornadas sobre Formación inicial del profesorado en Educación no sexista*. Madrid: Instituto de la Mujer.
- MATSUI, T., IKEDA, H. y OHNISHI, R. (1989): Relations of Sex-Typed Socializations to Career Self-Efficacy Expectations of College Students, *Journal of Vocational Behavior*, 35, 1-16.
- MESSING, K. (1983): The scientific mystique: Can a white lab coat guarantee purity in the search for knowledge about the nature of woman? En M. Lowe y R. Hubbard (Eds.): *Women's Nature: Rationalizations of Inequality*, (pp. 75-88). Nueva York: Pergamon Press.
- MINISTERIO DE EDUCACIÓN Y CIENCIA (1990): *I Jornadas de intercambio de experiencias de orientación para la igualdad de oportunidades entre ambos sexos*. Madrid: Autor.
- NUÑO, T. (1992): *Aproximación a la coeducación. Curso de sensibilización y orientación para desarrollar la coeducación en la escuela*. Comunicación presentada al Congreso Internacional sobre Investigación, Docencia y feminismo, San Sebastián, diciembre.
- PALMA, A. (1991): The Development of Non-Sexist Schools: Issues for Teacher Education, *Curriculum*, 12, 3, 165-171.

- PERDUE, V. P. y CONNOR, J. M. (1978): Patterns of touching between preschool children and male and female teachers, *Child Development*, 49, 1.258-1.262.
- PÉREZ-DELGADO, E., MESTRE, M. V. y GARCÍA-ROS, R. (1990): El efecto del género en el desarrollo del razonamiento sociomoral. Un estudio piloto sobre las diferencias mujer-varón a través del DIT en lengua castellana, *Revista de Psicología de la Educación*, 2, 5, 13-29.
- PIEL, E. (1991): Annual Review: Practice and Research in Career Counseling and Development, 1990, *The Career Development Quarterly*, 40, 2, 99-131.
- PIEL, E. (1993): The Gendered Context of life: Implications for Women's and Men's Career-Life Plans. *The Career Development Quarterly*, 41, 3, 227-237.
- PLATEAU, N. (1991): French-Speaking Belgium. En M. Wilson (Ed.): *Girl and Young Women in Education. A European Perspective*, (pp. 15-40). Oxford: Pergamon Press.
- POLYDORIDES, G. K. (1991): Greece. En M. Wilson (Ed.): *Girl and Young Women in Education. A European Perspective*, (pp. 91-114). Oxford: Pergamon Press.
- PREDIGER, D. J. y NOETH, R. J. (1979): Effectiveness of a brief counseling intervention in stimulating vocational exploration. *Journal of Vocational Behavior*, 14, 352-368.
- REA-POTEAT, M. B. y MARTIN, P. F. (1991): Taking your place: A Summer Program to Encourage Nontraditional Career Choices for adolescents girls. *The Career Development Quarterly*, 40, 2, 182-188.
- REAY, D. (1991): Intersections of Gender, Race and Class in the Primary School. *British Journal of Sociology of Education*, 12, 2, 163-182.
- REID, P. T. y COMAS-DÍAZ, L. (1990): Gender and ethnicity: Perspectives on dual status, *Sex Roles*, 22, 397-408.
- REINHARZ, S. (1992): *Feminist Methods in Social Research*. Nueva York: Oxford University Press.
- ROBINSON, K. H. (1992): Class-Room Discipline: power, resistance and gender. A look at teacher perspectives, *Gender and Education*, 4, 3, 273-287.
- ROBINSON, T. (1993): The Intersections of Gender, Class, Race, and Culture: On seeing clients whole, *Journal of Multicultural Counseling and Development* 21, 50-58.
- RUBIO, E. y MAÑERU, A. (1989): El género como categoría de análisis de la educación, *Revista de Educación*, 290, 7-20
- SADKER, M. P. y SADKER, D. M. (1985): Sexism in the schoolroom of the «80»s, *Psychology Today*, marzo, 54-57.
- SADKER, M. P.; SADKER, D. M. y KLEIN, S. (1991): The issue of gender in elementary and secondary education. En G. Grant (Ed.): *Review of research in education*, 17, 269-334.
- SCHAU, C. G. y SCOTT, K. P. (1984): Impact of gender characteristics of instructional materials: An integration of the research literature. *Journal of Educational Psychology*, 76, 183-193.
- SCHER, M. y GOOD, G. E. (1990): Gender and Counseling in the Twenty-First Century: What Does the Future Hold?, *Journal of Counseling and Development*, 68, 4, 388-391.
- SCHWEITZER, I. (1985): Auch heute noch: Sexismus in Fibeln?-Eine Analyse (Sexism in children's books? The debate continues): En R. Valtin y U. Warm (Eds.): *Frauen machen Schule (Woman make school)*: Frankfurt: Arbeitskreis Grundschule ev.
- SHILLING, C. (1991): Social Space, Gender Inequalities and Educational Differentiation, *British Journal of Sociology of Education*, 12, 1, 23-44.
- SIKES, P. J. (1991): «Nature Took its Course»? Student Teachers and Gender Awareness, *Gender and Education*, 3, 2, 145-162.
- SKAALVIK, E. M. y RANKIN, R. J. (1990): Math, Verbal, and General Academic Self-Concept: The Internal/External Frame of Reference Model and Gender Differences in Self-Concept Structure, *Journal of Educational Psychology*, 82, 3, 546-554.

- STICKEL, S. A. y BONETT, R. M. (1991): Gender differences in Career Self-Efficacy: Combining a Career with home and family, *Journal of College Student Development*, 32, 297-301.
- STOBART, G., ELWOOD, J. y QINLAN, M. (1992): Gender Bias in Examinations: how equal are the opportunities?, *British Educational Research Journal*, 18, 3, 261-276.
- STRACHAN, J. (1991): *Empowering Women for Educational Leadership*, M. Ed. Thesis, University of Waikato. Hamilton. Nueva Zelanda.
- STRACHAN, J. (1993): Including the Personal and the professional: researching women in educational leadership, *Gender and Education*, 5, 1, 71-80.
- STYFIELD, H. (1991): *The position of Women in Education, 1990*, Ministry of Education, Wellington, Nueva Zelanda.
- SUBIRATS, M. (1985): Mujer y educación: de la enseñanza segregada, a la coeducación. En Instituto de la Mujer (Ed.): *Primeras Jornadas sobre Mujer y Educación*, (pp. 9-21). Madrid: Instituto de la Mujer.
- SUBIRATS, M. (1987): La transmisión de estereotipos sexuales en el sistema escolar. En Instituto de la Mujer (Ed.): *La investigación en España sobre mujer y educación*, (pp. 20-24). Madrid: Instituto de la Mujer.
- SUBIRATS, M. y BRULLET, C. (1988): *Rosa y azul. La trasmisión de los géneros en la escuela mixta*. Madrid: Instituto de la Mujer.
- TEN DAM, G. T. M. y VOLMAN, M. M. L. (1991): Conceptualising Gender Differences in Educational Research: the case of the Netherlands, *British Journal of Sociology of Education*, 12, 3, 309-321.
- THOMSON, J. (1988): *Women in Education: management training needs*. Research Report, Education Department, University of Waikato, hamilton, Nueva Zelanda.
- TUSIN, L. (1991): The relationship of academic and social self-concepts with women's choice of teaching as a career: a longitudinal model, *Journal of Research and Development in Education*, 24, 2, 16-27.
- VASIL, L. (1992): Self-Efficacy Expectations and Causal Attributions for Achievement among Male and Female University Faculty, *Journal of Vocational Behavior*, 41, 259-269.
- WILSON, J. y DANIEL, R. (1981): The effects of a career-options workshop on social an vocational stereotypes, *Vocational Guidance Quarterly*, 29, 341-349.
- WILSON, L. W. y HANDLEY, H. M. (1986): Listening to adolescents: gender differences in science classroom interaction. En L. C. Wilkinson y C. B. Marrett (Eds.): *Gender Influences in Classroom Interaction*. Nueva York: Academic Press.
- WILSON, M. (1991) (Ed.): *Girls and Young Women in Education. A European Perspective*. Oxford: Pergamon Press.
- WORRALL, N. y TSARNA, H. (1987): Teachers' reported practices toward girls and boys in science and languages, *British Journal of Educational Psychology*, 57, 300-312.
- YANICO, B. J. (1978): Sex Bias in Career Information: Effects of language on attitudes, *Journal of Vocational Behavior*, 13, 26-34.
- YOUNG, J. W. (1991): Gender Bias in Predicting College Academic Performance: A New Approach Using Item Response Theory, *Journal of Educational Measurement*, 28, 1, 37-47.

Diferencias Individuales

- ALONSO TAPIA, J. (1991): *Motivación y aprendizaje en el aula: cómo enseñar a pensar*. Madrid. Santillana.
- ALONSO TAPIA, J. (1992): *¿Qué es lo mejor para motivar a mis alumnos? Análisis de los que los profesores saben, creen y hacen al respecto*. Madrid. Univ. Autónoma. Colección Cuadernos del ICE, nº 5.
- ALONSO TAPIA, J. (1993): Motivación y estrategias de aprendizaje: determinantes contextuales e influjo recíproco. En C. Monereo (Comp.): *Las estrategias de aprendizaje*. Barcelona: Domenech.

- ALONSO TAPIA, J. y MATEIOS, M. (1985): Comprensión lectora: modelos, entrenamiento y evaluación, *Infancia y aprendizaje*, 31-32, 5-19.
- ÁLVAREZ, V. y ADELMAN, H. (1986): Over-statements of self-evaluations by students with psychoeducational problems, *Journal of Learning Disabilities*, 18, 567-571.
- AMES, C. (1984): Competitive, cooperative and individualistic goal structures: a cognitive-motivational approach. En R. E. Ames y C. Ames (Ed.): *Research on motivation in education. Vol. 1: Student motivation*. Nueva York: Academic Press.
- AMES, C. (1992): Classroom: goals, structures and student motivation, *Journal of Educational Psychology*, 84, 3, 261-271.
- BANDURA, A. (1977): Self-efficacy: toward a unifying theory of behavior change, *Psychological Review*, 84, 191-215.
- BANDURA, A. (1982): Self-efficacy mechanism in human agency, *American Psychologist*, 37, 122-147.
- BANDURA, A. (1986): *Social foundations of thought and action: a social cognitive theory*. Englewood Cliffs, N. J. Prentice Halls.
- BARNETT, L. (1992): *El aprendizaje cooperativo y las estrategias sociales*. En C. Monereo (Comp) Las estrategias de aprendizaje. Barcelona: Domenech.
- BELTRÁN, J. (1992): *Programas para la mejora de la inteligencia*. Seminario Internacional sobre Programas para la mejora de la inteligencia. Madrid: Ponencia.
- BEREITER, C. (1980): Development in writing. En L. W. Gregg y E. R. Steinberg (Eds.): *Cognitive processes in writing*. Hillsdale, Nueva Jersey: LEA.
- BEREITER, C. y SCARDAMALIA, M. (1987): *The psychology of written composition*. Hillsdale, Nueva Jersey: LEA.
- BERNAD, J. A. (1993): Estrategias de aprendizaje y enseñanza: evaluación de una actividad compartida en la escuela. En C. Monereo (Comp): *Las estrategias de aprendizaje*. Barcelona: Domenech.
- BLUMENFELD, Ph. C. (1992): Classroom learning and motivation: clarifying and expanding goal theory, *Journal of Educational Psychology*, 84, 3, 272-281.
- BORKOWSKY, J. y otros (1990): Self-regulated cognition: interdependence of metacognition, attributions and self-esteem. En B. Jones y L. Idol: *Dimensions of thinking and cognitive instruction*, (pp. 381-413). Hillsdale, Nueva Jersey: Lawrence Erlbaum Asso.
- BRADLEY, C. A. (1985): The relationship between student's information-processing styles and Logo programming, *Journal of Educational Computing Research*, 1, 4, 427-433.
- BRANSFORD, J. D. y otros (1990): Teaching thinking and content knowledge: toward an integrated approach. En B. Jones y L. Idol: *Dimensions of thinking and cognitive instruction*. Hillsdale, Nueva Jersey: Lawrence Erlbaum Asso.
- BROPHY, J. (1983): Conceptualizing student motivation, *Educational Psychologist*, 18, 200-215.
- BROPHY, J. (1988): Research linking teacher behavior to student achievement: potential implications for instruction of chapter 1 students, *Educational Psychologist*, 23, 235-286.
- BURWELL, L. B. (1991): The interaction of learning styles with learner control treatments in an interactive videodisc lesson, *Educational Technology*, marzo, 37-43.
- CABERO, J. (1989): *Tecnología educativa: utilización didáctica del vídeo*. Barcelona: P.P.U.
- CABERO, J. y SALAS, M. (1990): Estilos cognitivos e investigación en medios de enseñanza, *Campo Abierto*, 7, 26-40.
- CALVO, A. (1990): *Actitudes autoritarias y maquiavélicas en estudiantes de enseñanza media*. Tesis Doctoral. Madrid: Universidad Complutense.
- CALVO, A. (1992): El aprendizaje cooperativo en el aula y la modificación de actitudes autoritarias, *Cuadernos de Pedagogía*, 21-28.
- CARRETERO, M. (1982): Dependencia-Independencia de campo, *Infancia y Aprendizaje*, 18, 65-83.

- CARRETERO, M. y PALACIOS, J. (1982): El desarrollo de los estilos cognitivos: breve presentación de un amplio tema, *Infancia y aprendizaje*, 17, 19-29.
- CARRIEDO, N. (1993): Enseñar a comprender las ideas principales en el aula. En C. Monereo (Comp.): *Las estrategias de aprendizaje*. Barcelona: Domenech.
- CASTELLÓ, M. (1993): Las estrategias de aprendizaje y la composición escrita. En C. Monereo (Comp.): *Las estrategias de aprendizaje*. Barcelona. Domenech.
- CLEMENTS, D. H. y GULLO, D. F. (1984): Effects of computer programing on young children cognition, *Journal of Educational Psychology*, 76, 1.051-1.058.
- CORNO, L. y SNOW, R. E. (1986): Adapting teaching to individual differences among learners. En C. M. Wittrock (Ed.): *Handbook of research on teaching*. London: MacMillan.
- DANSERENAU, D. F. (1985): Learning strategy research. En J. Segal, S. Chipman y R. Glaser (Eds.): *Thinking and learnig skills, 1: Relating instruction to research*. Hillsdale, Nueva Jersey: Lawrence Erlbaum Ass.
- DE BONO, E. (1983): The cognitive research trust (CORT) thinking program. En W. Maxwell (Ed.): *Thinking: the expanding frontier*. Philadelphia: The Franklin Institute.
- DECHARMS, R. (1976): Enhacing motivation:change in the classroom. Nueva York: Irvington.
- DECI, E. L. (1975): *Intrinsic motivation*. New York. Plenum Press.
- DECI, E. L. y RYAN, R. M. (1985): Intrinsic motivation and self-determination in human behavior. New York: Plenum Press.
- DUFFY, G. y ROEHLER, L. (1989): *Improving classroom reading instruction: a decision-making approach*. N. York: Random House.
- DUNN, R. y DUNN, K. (1975): *Educator's self-teaching guide to individualizing instructional programs*. West Nyack, Nueva York: Parker Pu. Co.
- DUNN, R., DUNN, K. Y PRICE, G. E. (1977): *Identifying individual learning styles and the instructional methods and/or resourdes to wich they respond*. Documento presentado en la reunión anual de la Asociación de Investigación Educativa. Nueva York, abril.
- DUNN, R. y DUNN, K. (1984): *La enseñanza y el estilo individual de aprendizaje*. Madrid: Anaya.
- FEUERSTEIN, R. (1980): Instrumental Enrichment. Baltimore: University Park Press.
- FOUX, B. y ROUTH, D. K. (1984): Phonemic analysis and synthesis as words attack skills: revisited, *Journal of Educational Psychology*, 76, 6, 1.059-1.064.
- FULLER, R. G. (1980): Active learning based upon the work of Piaget. En R. G.Fuller y otros (Comp.): *Piagetian programs in higher education*. Lincoln, N.E.: ADAPT Program.
- GARNER, R. (1990): When children and adults do not use learning strategies: toward a theory of settings, *Review of Educational research.*, 60, 4, 517-529.
- GLASER, R. (1984): Thinking and learning: the role of knowledge, *American Psychologist*, 39, 2,93-104.
- GOTTFRIED, A. E. (1990): Academic intrinsic motivation in young elementary school children, *Journal Educational Psychology*, 82, 525-538.
- GRAHAM, S. y GOLAN, S. (1991): Motivational influences on cognition: Task involvement, ego involvement, and depth of information processing, *Journal of Educational Psychology*, 83, 187-194.
- GRAHAM, S. y HARRIS, K. R. (1989): Components analysis of cognitive strategy instruction:effects on learning disabled students compositions and self-efficacy, *Journal of educational Psychology*, 81, 353-361.
- HANNAFIN, M. J., PHILIPS, T. L. y TRIPP, S. D. (1986): The effects of orienting activities rehearsal and processing time during interactive video, *Journal of Computer-Based Instruction*, 13, 4, 134-139.
- HARVARD UNIVERSITY (1983): *Project Intelligence: the development of procedures to enhance thinking skills*. Informe Final.Venezuela.
- HOOPER, S. (1992): Cooperative learnig and computer-based instruction, *Educational Technology resear- ch and development*. 40, 3, 21-53.

- JOHNSON, D. W. y JOHNSON, R. T. (1990): Cooperative learning and achievement. En S. Sharan. *Cooperative learning*, (pp. 23-35). Nueva York: Praeger.
- KAGAN, N. (1981): Las implicaciones de los estilos cognitivos en la educación. En G. S. Lesser (Ed.): *La psicología en la práctica de la educación*, (pp. 303-366). Méjico: Trillas
- KELLER, J. (1987): Development and use of the ARCS model of instructional design, *Journal of Instructional Development*, 10, 2-10.
- KING, A. (1990): Enhancing peer interaction and learning in the classroom through reciprocal questioning, *American Educational Research Journal*, 27, 664-687.
- LIPMAN, M. y otros (1980): *Philosophy in the classroom*. Filadelfia: Temple University Press.
- MARSHALL, H. H. (1987): Motivational strategies of three fifth-grade teachers, *The Elementary School Journal*, 88, 135-150.
- MAYER, R. E. (1987): *Pensamiento, resolución de problemas y cognición*. Madrid: Paidós/MEC.
- MCROBBIE, C. J. (1991): Cognitive styles and cognitive structure, *Science education*, 75, 2, 231-242.
- MILLER, A. y HOM, H. L. (1990): Influence of extrinsic and ego incentive value on persistence after failure and continuing motivation, *Journal of Educational Psychology*, 82, 539-545.
- MILLER, J. E. y EMIWOVICH, C. (1988): Effects of Logo and CAI on black first grades, achievement, reflexivity, and self-esteem, *Elementary School Journal*, 88, 5, 473-487.
- MONEREO, C. (1993): *Las estrategias de aprendizaje*. Barcelona: Domenech.
- MULTON, K., BROWN, S. y LENT, R. (1991): Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation, *Journal of Counseling Psychology*, 38, 30-38.
- NICHOLLS, J. G. (1984): Achievement motivation: Conceptions of ability, subjective experience, task choice, and performance, *Psychological Review*, 91, 328-346.
- NICHOLLS, J. G. (1989): *The competitive ethos and democratic education*. Cambridge, M.A.: Harvard University Press.
- NICKERSON, R. S. (1987): *Enseñar a pensar*. Madrid. Paidós-MEC.
- NEWBY, T. J. (1991): Classroom motivation: strategies of first-year teachers, *Journal of Educational Psychology*, 83, 195-200.
- NISBET, J. (1991): Investigación reciente en estrategias de estudio y el enseñar a pensar. En C. Monereo (Comp.): *Enseñar a pensar a través del currículum escolar*, (pp. 11-19). Barcelona: Casallas/COMAP.
- NOLEN, S. B. (1988): Reasons for studying: Motivational orientations and study strategies, *Cognition and Instruction*, 5, 269-287.
- O'DONELL, A., DANSEREAU, D. F. y ROCKLIN, T. R. (1991): Individual differences in the cooperative learning of concrete procedures, *Learning and Individual differences*, 3, 2, 149-162.
- ORJALES, I. y POLAINO-LORENTE, A. (1992): Estilos cognitivos e hiperactividad infantil: los constructos dependencia-independencia de campo perceptivo e impulsividad-reflexividad, *Bordón*, 44, 4, 421-430
- PARDO MERINO, A. y ALONSO TAPIA, J. (1990): *Motivar en el aula*. Madrid: Universidad Autónoma. Colección Estudios.
- PALINCASAR, A. y otros (1988): Collaborative research and development of reciprocal teaching, *Educational leadership*, 46, 37-41.
- POST, P. (1987): The effects of field independence-field dependence on computer assisted instruction achievement, *Journal of industrial Teacher Education*, 25, 1, 60-67.
- PRESSLEY, M. y otros (1990): A primer research on cognitive strategy instruction: the important issues and how to address them, *Educational Psychology Review*, 2, 1-57.
- PRICE, G. E. (1984): Investigación sobre los estilos de aprendizaje y sugerencias a los profesores en este terreno. En R. Dunn y K. Dunn: *La enseñanza y el estilo individual de aprendizaje*. Madrid: Anaya.
- POZO, J. I. y otros (1993): *Las estrategias de aprendizaje como contenidos procedimentales*. Madrid: CIDE.

- RIDING, R. y SADLER-SMITH, E. (1992): Type of instructional material, cognitive style and learning performance, *Educational studies*, 18, 3, 323-340.
- ROMAN, J. M. (1993): Entrenamiento en estrategias de aprendizaje: secuencias, principios y validación. En C. Monereo (Comp.): *Las estrategias de aprendizaje*. Barcelona: Domenech.
- ROWLAND, y otros (1988): Matching mode of CAI to cognitive style: an exploratory study, *Journal of Computers in Mathematics and Science Teaching*, 7, 4, 36-40.
- SAWYER, R. J. GRAHAM, S. y HARRIS, K. R. (1992): Direct teaching, strategy instruction, and strategy instruction with explicit self-regulation: effects on the composition skills and self-efficacy of students with learning disabilities, *Journal of Educational Psychology*, 84, 340-352.
- SCARDAMALIA, M. y otros (1979): The little red writing book: a source book of consequential writing activities. Ontario, Canadá: Pedagogy of Writing Project, O.I.S.E.
- SCHMECK, R. (1988): Individual differences and learning strategies. En C. E. Weinstein y otros (Eds.): *Learning and study strategies: issues in assessment, instruction and evaluation*. Londres: Academic Press.
- SELMES, I. (1986): Approaches to normal learning tasks adopted by senior secondary school pupils, *British Educational Research Journal*, 12, 1, 14-27.
- SELMES, I. (1988): *La mejora de las habilidades para el estudio*. Madrid: Paidós-MEC.
- SLAVIN, R. E. (1990): *Cooperative learning: theory, research and practice*. Englewood Cliffs, Nueva Jersey: Prentice Hall.
- STEVENS, R. y otros (1987): Cooperative integrated reading and composition: two field experiments, *Reading Research Quarterly*, 22, 443-454.
- STEVENS, R. J., SLAVIN, R. E., y FARNISH, A. M. (1989): *A cooperative learning approach to elementary reading and writing instruction: long-term effects*. Baltimore: Johns Hopkins University, Center for Research on Elementary and Middle schools.
- STEVEN, R. J., SLAVIN, R. E., y FARNISH, A. M. (1991): The effects of cooperative learning and direct instruction in reading comprehension strategies on main idea identification, *Journal of Educational Psychology*, 83, 1, 8-16.
- SOLÉ, I. (1993): Las estrategias de lectura como estrategias de aprendizaje. En C. Monereo (Comp.): *Las estrategias de aprendizaje*. Barcelona: Domenech.
- VIDAL-ABARCA, E. y GILABERT, R. (1989): Estrategias para la comprensión y el recuerdo de la información más importante en textos expositivos: un estudio experimental, *Revista de psicología de la Educación*, 2, 1-20.
- VIDAL-ABARCA, E. y GILABERT, R. (1991): *Comprender para aprender: un programa para mejorar la comprensión y el aprendizaje de textos*. Madrid: CEPE.
- WATSON, J. (1990 y 1991): Cooperative learning and computers: one way to address student differences, *The Computing Teacher*. Dic-Enero, 9-13.
- WEINER, B. (1986): *An attributional theory of motivation and emotion*. New York: Springer-Verlag.
- WEINER, B. (1990): History of motivational research in education, *Journal of Educational Psychology*, 82, 616-622.
- WEINSTEIN, C. (1988): *Learning and study strategies*. S. Diego: Academic Press.
- WITKIN, H. A. (1962): *Psychological differentiation: studies of development*. Nueva York: Wiley.

PONENCIA II

PERSPECTIVAS METODOLÓGICAS ACTUALES DE LA EVALUACIÓN DE PROGRAMAS EN EL ÁMBITO EDUCATIVO

por

F. Javier Tejedor, Ana García-Valcárcel y M^a José Rodríguez Conde
Universidad de Salamanca

1. CONCEPTUALIZACIÓN DE LA EVALUACIÓN DE PROGRAMAS

La «evaluación de programas» surge en el ámbito educativo en el entorno de los años setenta con el objetivo fundamental de valorar la eficacia de los programas de acción social que se llevaban a cabo, ya que no había constancia alguna de la consecución de los objetivos previstos. La evaluación de programas de actuación pública o de intereses sociales supone, sobre todo a medida que el programa es ambicioso, analizar realidades complejas, dinámicas y en las que se desarrollan procesos de cambio de sistemas interdependientes (Colás, 1993, p. 10).

La evaluación de programas, dicen Cook y Shadish, es un modelo racional de solución de problemas de carácter psicológico, social, educativo,..., que requiere que: (1) los problemas se definan primero con claridad y después... (2) se genere una amplia variedad de soluciones potenciales... (3) algunas de las cuales se implementan y... (4) finalmente se evalúan. (5) El conocimiento de las soluciones positivas se presentan a los responsables políticos quienes... (6) lo utilizarán para tomar las decisiones políticas que correspondan (Cook y Shadish, 1986, p. 197).

La evaluación de un programa se planifica para tomar decisiones, por tanto los datos que se busquen deben facilitar dicha toma de decisiones. Debe ser un proceso cíclico y continuo, cuya puesta en marcha responda a una programación sistemática, para valorar si un programa consiguió o no alcanzar los objetivos para cuya consecución fue ideado. Este proceso consta de cuatro actividades básicas (Ato y otros, 1989):

- identificar los objetivos del programa
- transformar los objetivos en variables observables (indicadores)

- recoger datos empíricos relativos a las variables identificadas
- comprobar los objetivos logrados con los datos empíricos obtenidos, valorando el éxito o fracaso en su consecución y el balance económico del binomio coste/eficacia.

La finalidad de la evaluación es, pues, emitir un juicio valorativo sobre el programa y determinar en qué medida logra los objetivos marcados. Trata de maximizar la eficiencia y la eficacia de las acciones dirigidas a modificar segmentos de la realidad.

Hay que destacar también que la evaluación de programas pretende explicar los hallazgos encontrados (relación entre variables, incidencia de unas variables en otras,...) y no únicamente describir los resultados del proceso de medición.

La evaluación de programas ha ido perfilándose con una metodología concreta en los últimos quince años. Esta metodología queda delimitada por: una terminología propia, un conjunto de herramientas conceptuales y analíticas específicas, y unos procesos y fases también específicos. El conjunto define un campo floreciente en el que realmente lo específico es la mezcla y no las partes constitutivas.

Para Alvira (1991), algunas circunstancias que han contribuido a conformar este campo de actuación serían:

- a) Mayor interrelación entre las actividades de evaluación y programación. Si en una perspectiva clásica la evaluación se realiza después de las actividades de planificación, programación y desarrollo del programa, en la actualidad ambas actividades caminan de forma paralela. Es decir, frente al esquema de actuación tradicional se plantea una estrategia con mayor interrelación entre las partes, sin que ello suponga una pérdida de especificidad de cada uno de los componentes (cuadro 1):
- b) La evaluación de programas, desde la perspectiva metodológica que es la que aquí tratamos, es investigación evaluativa; es decir, es investigación (con todos los considerandos epistemológicos, metodológicos y técnicos que ello entraña), pero es más que investigación, ya que se entiende como un proceso cuya finalidad es aplicar procedimientos científicos para acumular evidencia válida y fiable sobre la manera y grado en que un conjunto de actividades producen resultados o efectos concretos. Intenta determinar el valor de la acción educativa, buscar soluciones a problemas concretos o mejorar las estrategias de actuación.
- c) En la evaluación de programas se produce una mayor implicación de las partes interesadas, lo que no necesariamente implica mayor aproximación a estrategias cualitativas (observación participante, investigación acción,...); lo que sí implica necesariamente es un mayor esfuerzo por atender las demandas de quienes patrocinan el trabajo y de los potenciales beneficiarios del mismo.
- d) En la evaluación de programas pueden integrarse estrategias de investigación muy diversas, tanto cuantitativas como cualitativas. Es un campo de trabajo que posibilita la utilización simultánea y sucesiva de diferentes metodologías para el tratamiento de los temas objeto de estudio e intervención. Será la

CUADRO 1
PERSPECTIVA ACTUAL EN EVALUACIÓN DE PROGRAMAS

Estrategia tradicional	Estrategia actual
1) Identificación de problemas y/o necesidades	Evaluación de necesidades
2) Programación y priorización de objetivos	Evaluación conceptual
3) Diseño del programa de intervención elegido	Evaluación del modelo
4) Ejecución del programa	Evaluac. de implementación Evaluación de la cobertura Evaluación del proceso
5) Evaluación	Evaluación de resultados

naturaleza específica del programa y la pericia del evaluador quienes a la postre determinarán la utilización de una o varias opciones metodológicas.

La evaluación de programas en el ámbito español, como señalan Colás y Rebollo (1993, p. 9), se caracteriza por las escasas realizaciones sistematizadas y el predominio de determinados enfoques.

A nivel internacional, la situación actual se define por la utilización de múltiples modelos que se aplican a distintos contenidos (necesidades, diagnóstico, productos, impactos,...), la multidisciplinariedad (análisis etnográfico y estadístico), multidimensionalidad en objetivos, en perspectivas y en niveles de valoración (validez, credibilidad, costes, rentabilidad, aplicabilidad). En síntesis, múltiples metodologías, múltiples impactos, múltiples formatos de informe,...Es una multiplicidad de lo múltiple (Colás y Rebollo, 1993, citando a Conner).

Una consecuencia de este panorama es sin duda la dificultad de conceptualizar la «evaluación de programas», encontrándonos con múltiples definiciones en las que cada autor destaca los elementos que más le interesan. En la búsqueda de patrones de similitud nos encontramos dos tipos básicos de apreciaciones: los que hacen hincapié en la dimensión técnica y los que ponen el énfasis en los considerandos ideológicos, aunque ni siquiera estas caracterizaciones extremas pueden considerarse tan distintas ya que realmente no son opuestas ni contradictorias, resultando en el fondo complementarias por estar destacando cada una de ellas momentos distintos del proceso: selección del contenido del programa, toma de decisiones, implicación

de los participantes, en un caso, y recogida y análisis de información orientada a la toma de decisiones, en otro.

Los considerandos básicos que caracterizan la evaluación de programas, en cualquiera de sus modalidades, serían, en opinión de Colás (1993) y Gairín (1991) que nosotros reinterpretemos, los siguientes:

- La evaluación de programas no debe entenderse como un proceso uniforme con un modelo único de evaluación o con una única metodología.
- La evaluación de programas se ajustará inicialmente a la idea que se tenga de evaluación. La concepción más amplia es aquella que entiende que evaluación es el conocimiento empírico del mérito o valor de las actividades y objetivos de un programa.
- Irá referida a un programa o conjunto sistemático de actuaciones que se ponen en marcha para alcanzar unos objetivos.
- Junto a la función educativa, el programa puede estar cumpliendo otras funciones de carácter ideológico y político.
- Planificación y evaluación deben pensarse simultáneamente.
- La evaluación de necesidades es una tarea previa a la planificación de un programa.
- Los modelos de evaluación deben ser al mismo tiempo modelo de análisis y de toma de decisiones.
- La planificación-evaluación del programa debe ser detallada si queremos facilitar las tareas posteriores de control.
- El evaluador debe estar familiarizado con el objeto de la evaluación (contacto y negociación previa).
- La participación de los interesados es deseable ya que aporta validez y es garantía de utilidad y aplicabilidad. Aunque en algunos diseños no se incluye de forma explícita o se duda de su utilidad, creemos que el problema actualmente no se plantea en términos de si admitir o no la participación de los interesados (gestores y destinatarios) en el proceso sino el delimitar la forma y el grado en que debe producirse (en algunos modelos, la participación es una fase del proceso).
- Los criterios e indicadores son básicos en la evaluación así como el establecimiento de estándares o grados aceptables de éxito.
- La evaluación se entiende como proceso, confiriéndole un carácter dinámico. El concepto de evaluación concreto que se adopte marcará la elección de posibles modelos teóricos y enfoques metodológicos a seguir, así como otros aspectos a considerar (tema, actividades, rol del evaluador, audiencia, informes,...).
- Deben aplicarse procedimientos científicos, que incluyan estrategias de diseño, recogida y análisis de información rigurosos y sistematizados, lo que no excluye en modo alguno la diversidad de enfoques metodológicos (técnicas de recogida y análisis de datos y estrategias de toma de decisiones). Tampoco el carácter científico del proceso evaluador debe hacernos concebir dicho

proceso como una actividad exclusivamente teórica, mera aplicación de los procedimientos científicos.

- Actualmente el énfasis de la tarea evaluativa se pone en la emisión del juicio de valor, a partir de una adecuada recogida y análisis de la información pertinente.
- Deben superarse las opciones excluyentes de metodologías específicas; la evaluación de programas permite utilizar de forma conjunta y coordinada metodologías diferentes.
- No todos los diseños son adecuados. Deben responder a criterios de validez y utilidad.
- Debe tener como proyección la toma de decisiones, lo que implica otorgarle una utilidad a la evaluación.
- Se remarca el carácter contextualizado, en el sentido de que ha de estar condicionada por las circunstancias y ha de ajustarse a las condiciones reales en las que se aplica el programa objeto de evaluación.
- La tipología de la evaluación depende del momento o fase del programa que quiera analizarse: de necesidades, de proceso, de producto,...
- Hay que prestar especial atención a los efectos no previstos, ya que pueden deberse tanto a la potencialidad del programa como a su incorrecta aplicación.
- El éxito de un programa no se mide únicamente por la consecución de objetivos ya que hay programas que no los formulan o que se plantean desde otras perspectivas menos funcionales.

2. BREVE PERSPECTIVA HISTÓRICA

Para algunos autores el comienzo de la Evaluación de Programas, tal y como hoy se entiende, se sitúa en la década de los años cincuenta en el entorno de las actividades de evaluación de objetivos educativos llevadas a cabo por Tyler.

En los planteamientos de Tyler la evaluación debe cumplir el papel de verificar la consecución de los objetivos operativos propuestos previamente a la realización de la actividad (bajo el supuesto de que esta realización posibilita la consecución de dichos objetivos). Cuanto más claramente estén explicitados los objetivos más fácil resultará la verificación o no de su consecución. La diferencia entre las medidas pre y post programa será un indicador preciso de sus efectos. Los esfuerzos metodológicos por garantizar la exactitud de esas medidas (y por tanto del juicio emitido) vinieron dados por Campbell y Stanley con la publicación en 1973 de su conocida obra sobre los diseños de investigación. Las pautas metodológicas contenidas en esa obra orientan todavía hoy la actividad de investigación evaluativa, aunque en el entorno de los años ochenta, con el cuestionamiento de la viabilidad de los estudios experimentales en el ámbito de las ciencias sociales, se abren nuevas perspectivas que suponen la incorporación de nuevas estrategias metodológicas, tanto cuantitativas (diseños correlacionales, series, técnicas de clasificación,...), como cualitativas.

Presentamos, de forma esquemática, un breve panorama de la evolución histórica de las prácticas evaluativas.

1) Antecedentes: En torno a los años 20 (evaluación de 1ª generación); corriente positivista y empiricista; se centra en la medición de características individuales (testing).

2) Consolidación: Años 30-50 (evaluación de 2ª generación); referencia a los objetivos (Tyler); se pretende conseguir la descripción de la medida.

3) Auge: Años 55-70 (evaluación de 3ª generación); evolución de la etapa de la inocencia al realismo; se pretende medir, describir e interpretar; predomina la evaluación como investigación empírica positiva, que utiliza procedimientos estandarizados para la recogida y análisis de datos orientados a la toma de decisiones.

4) Actualidad: A partir de los años 70; investigación evaluativa (diversidad de paradigmas); coincidiendo con la aparición de las prácticas cualitativas, la evaluación de programas incorpora modelos que se centran en los participantes; los procedimientos de investigación incorporados provienen del campo de la investigación cualitativa (naturalista, etnográfica,...); a partir de 1980 (evaluación de 4ª generación) se habla del modelo crítico de evaluación, sin una personalidad muy reconocida ya que la mayoría de los autores lo consideran una modalidad de la perspectiva cualitativa; actualmente, se presenta la evaluación como negociación (Guba y Lincoln).

Los ochenta, en opinión de Alvira (1991), se caracterizan por:

- Una expansión de la evaluación de programas a los países occidentales a la par que se restringe el dinero para éstas en Estados Unidos.
- La aparición de una pluralidad de perspectivas, exageradamente denominadas modelos, y de nuevos métodos.
- Un énfasis creciente en lograr que las evaluaciones sean útiles.
- Ampliación del tipo de preguntas a las que debe dar respuestas una evaluación, rebasando con creces la evaluación de resultados que había sido predominante anteriormente.

Hoy se vive en la metodología de la evaluación una situación plural, compleja y que tiende al realismo en mayor grado que anteriormente; las nuevas perspectivas no sustituyen a las anteriores sino que las complementan. Los cambios habidos en estos últimos años en la conceptualización de la evaluación de programas quedan recogidos en el cambio recogido en las dos posturas siguientes: la postura más tradicional que entendía la evaluación como el empleo de métodos científicos para medir el desarrollo y los resultados de programas que luego se utilizarán para la toma de decisiones, y la postura más actual que entiende la evaluación como el conocimiento empírico del mérito o valor de las actividades y objetivos de un programa.

3. LA INVESTIGACIÓN EVALUATIVA Y LOS PROCESOS DE INTERVENCIÓN

La investigación evaluativa es una manera de aumentar la racionalidad de las decisiones en relación con la puesta en marcha de programas sociales, evitando una excesiva dependencia de criterios políticos. Los objetivos de la investigación evaluativa suelen ser de dos tipos: elaborar juicios sobre la eficacia de los programas y ayudar en las decisiones a los responsables.

La investigación evaluativa es un modo de investigación que implica un proceso riguroso, controlado y sistemático de recogida y análisis de información fiable y válida para tomar decisiones sobre un programa educativo. Vemos que aparecen importantes puntos de contacto respecto a lo que entendemos por «investigación»: proceso sistemático, riguroso, controlado; utilización de técnicas de medición y análisis de datos (la investigación se distingue de la evaluación por los objetivos y no por los métodos).

Sin embargo, creemos que conviene diferenciar ambas tareas. Las características peculiares de la investigación evaluativa que señala De la Orden (1985 b) son:

- a) Los juicios de valor acompañan al desarrollo de todo el programa: desde la selección del problema a estudio hasta en la aplicación de metodologías concretas.
- b) Resulta difícil (a veces inadecuado) formular hipótesis precisas.
- c) La replicación es prácticamente imposible dadas las características casi irrepetibles de la ejecución de un programa de intervención.
- d) La recogida de datos está condicionada por la viabilidad del proceso.
- e) Multitud de variables relevantes sólo son susceptibles de un control superficial limitándose por tanto la utilización de diseños de carácter más experimental.
- f) Las decisiones importantes en relación con la paralización, sustitución o repetición del programa no son competencia del evaluador.
- g) El informe debe adaptarse a las exigencias de quien lo financia.

En el cuadro 2 se recogen, para algunos de los tópicos principales, las diferencias comentadas (y algunas otras) entre investigación (en términos generales) e investigación evaluativa.

Una característica importante de la investigación evaluativa es que se lleva a cabo en un proceso de intervención, en la misma acción, que es lo realmente importante. La investigación, señalaba Weis (1982), es una prioridad secundaria, lo que debe llevarnos a pensar que la evaluación/investigación tiene que adaptarse al contexto del programa y perturbar lo menos posible su desarrollo.

Esta aproximación de la investigación a la práctica educativa, ¿la hace ser más útil? Puesto que la Evaluación de Programas está siendo planteada, desde una perspectiva metodológica, en términos de investigación evaluativa, al plantearnos su utilidad nos estamos realmente preguntando por la utilidad de los resultados de la investigación. La respuesta, en todo caso, es compleja. En numerosas ocasiones se

CUADRO 2
INVESTIGACIONES VS INSVESTIGACIÓN EVALUATIVA

<u>Tópicos</u>	<u>Investigación</u>	<u>Investigación evaluativa</u>
Propósito	incrementar el conocimiento sacar conclusiones	modificar la realidad tomar decisiones
generalización de resultados	a otros ámbitos (validez externa)	limitado a grupo de referencia
juicios de valor	conocimiento objetivo (la verdad)	asignar criterios de valor
papel investigador	autónomo	dependiente del gestor
temas	interés personal y/o científico	interés del gestor
metodología	básicamente cuantitativa	diversas
diseño	preestablecido	preestablecido pero flexible
hipótesis	explícitas	no explícitas
replicación	exigible	no exigible
datos	a partir del problema	a partir de la viabilidad
control y manipulación	deseable (exigible en la básica)	deseable, pero muy difícil
aleatorización	básica	no procede
criterios de validez	interna, externa y de constructo	utilidad credibilidad
interpretación de datos	de acuerdo a reglas definidas	valoración por los responsables del programa
informe	adaptado a comunidad científica	adaptado a usuario y gestores

han puesto de manifiesto las limitaciones que concurren en la investigación educativa para proyectar los resultados a la práctica cotidiana.

El proceso de evaluación de programas debe concebirse como una auténtica estrategia de investigación sobre los procesos educativos, en cuyos resultados deberían basarse las pautas sugeridas para orientar los procesos de intervención. Investigación e intervención se apoyan mutuamente. La intervención sacará partido de la investigación en la medida en que ésta se extienda a la solución de problemas prácticos. La investigación activa se asimila, enriqueciéndola, a la noción de investigación aplicada.

El método, como forma de la actividad humana, puede adoptar dos modalidades básicas, según que dicha actividad sea de carácter humano interno (métodos de pensamiento) o de carácter externo (métodos de acción). Los métodos de acción pueden a su vez concebirse como: a) técnicas o métodos de investigación, orientados al conocimiento de la realidad observada y b) técnicas o métodos de actuación dirigidas a transformar la realidad.

Toda situación educativa es susceptible de este doble tratamiento. Por una parte, existe la necesidad de «conocerla» y de «explicarla». Por otra, es preciso «comprenderla» y «mejorarla». Se requiere que la investigación alcance un nivel de integración suficiente, en este caso, entre el pensamiento y la decisión.

Entre el terreno del pensamiento teórico, donde sólo interesa el saber y en el que sólo el saber tiene valor propio, y el de la acción informada, donde lo útil y lo realizable van en primer lugar, hay una vasta zona en las ciencias del hombre donde las preocupaciones teóricas y utilitarias se mezclan en distinto grado: la preocupación por el saber se une de alguna manera a la preocupación por el actuar. Este campo de estudio, que algunos denominan «investigación orientada», está a la base de toda investigación evaluativa.

La investigación orientada/evaluativa nace como respuesta a necesidades sociales y se elabora en función de tales necesidades; se realiza con objeto de llegar a la solución de los problemas que exige una acción informada científicamente. Habrá de tener necesariamente un carácter interdisciplinar y viene determinada por los objetivos, los resultados y las motivaciones.

En estos momentos, lo que más nos podría interesar sería analizar los cauces que sigue la investigación orientada para contribuir a la utilización práctica de los resultados obtenidos, es decir, la contribución a la toma de decisiones, y a la inspiración de la acción educativa.

Pienso que la evaluación de programas, como técnica de investigación sobre los procesos de intervención socioeducativos, es una estrategia que contempla en su realización un potencial riquísimo para contribuir a la mejora de la calidad de la acción educativa pues tiene todos los requisitos exigibles a la investigación educativa útil.

No obstante, habría que preguntarse si siguen hoy siendo válidas las pesimistas reflexiones que Weis (1982) exponía hace unos años al valorar como no demasiado fructífera la breve historia de la investigación evaluativa; es más, decía, algunos

autores la consideran decepcionante. Pocos ejemplos de aportaciones importantes a las políticas y/o los programas pueden citarse. Ello se debe, en parte, a las notables resistencias de las organizaciones a ser evaluadas, a que sus programas sean evaluados, a ofrecer la información que se precisa en la forma en que se necesita. Otra parte de la culpa estriba en la manera en que a veces está estructurada, dotada de personal y ejecutada la misma evaluación. En ocasiones hay incoherencias entre los fines propuestos en la evaluación y el tipo de estudios o programas evaluados. Gran parte de la decepción que venimos comentando pueden deberse al carácter poco realista de las expectativas que se plantean, a la necesidad de continuidad en los estudios, a los esfuerzos por presentar adecuadamente el programa a los usuarios,...

4. MODELOS DE EVALUACIÓN. TIPOS Y PAUTAS PARA SU SELECCIÓN

En estos últimos años ha proliferado, quizá en exceso, la formalización de modelos de evaluación, lo que ha supuesto un cierto confusionismo para el usuario. Algunos autores opinan (Alvira, 1991) que en la mayoría de los casos no se trata de modelos en sentido estricto sino de diferentes aproximaciones a unas pocas variantes básicas de la metodología de la evaluación.

Recordando que la evaluación de programas requiere la emisión de juicios de valor (y por tanto el establecimiento previo de criterios para la realización de esos juicios) podría entenderse que las diferentes perspectivas difieren simplemente en los criterios de valor utilizados, que vendrán determinados por los propios objetivos de la evaluación.

Algunos de estos modelos básicos son los conocidos modelos de Tyler, Stake, Schuman, Stufflebeam, Scriven, ... Dado el carácter de este escrito no vamos a entrar en su descripción; en la bibliografía se ofrecen referencias para que el lector interesado pueda profundizar en el estudio de cada uno de ellos.

Los elementos básicos que componen un modelo de evaluación son, en opinión de Colás (1992), los siguientes: finalidad científica y/o política, contenido, toma de decisiones, ámbito o unidad de evaluación, rol del evaluador, enfoque y proceso metodológico. Como vemos, en la evaluación de programas el modelo es una síntesis entre ideología y metodología: se fundamenta ideológicamente y se concreta después es una propuesta metodológica específica, aceptándose actualmente cinco modalidades principales: experimentalista, ecléctica, descriptiva, coste-beneficio y crítica (Talmage, 1982). Y siempre aceptando mejor a quienes defienden la complementariedad entre los distintos métodos que a quienes la rechazan.

Creo que la opinión mayoritaria actual sobre esta temática es clara: actitud de búsqueda de la complementariedad de los diferentes enfoques, evitando la oposición, y sobre todo la descalificación, a posturas diferentes a las personales. Esta posición ecléctica, vinculada a mayores dosis de pragmatismo, entiende que la relación entre paradigmas y métodos es de tipo instrumental, siendo la metodología una simple colección de técnicas y procedimientos (Cook y Reichardt, 1986). Cada día parece más evidente que un determinado problema, el análisis de una determi-

nada realidad, puede ser realizado desde distintas perspectivas; lo realmente importante será la coherencia entre los objetivos del estudio, el tipo de datos y los análisis que se lleven a cabo. Con estudios bien hechos —sean del tipo que sean— resultará más fácil detectar similitudes interpretativas a través de los procedimientos adecuados (triangulación, matriz multirrasgo-multimétodo, confrontación de resultados,...).

Se han utilizado distintos criterios parciales (contenidos y efectos de la evaluación, enfoque metodológico,...) para presentar diferentes clasificaciones de los modelos de evaluación. En algún caso, se mezclan en esos criterios aspectos teóricos y metodológicos lo que proporciona esquemas más coherentes. Seleccionamos para presentar aquí las propuestas de clasificación de Popham y Rebollo.

Popham (1980) nos sugiere la siguiente clasificación de los modelos:

a) *Modelos de concreción de metas:*

— *Modelo de Tyler*, caracterizado por:

- establecer las metas y objetivos
- ordenar los objetivos en amplias clasificaciones
- definir los objetivos en términos operativos
- establecer en qué condiciones puede demostrarse la consecución de objetivos
- explicar los propósitos de la evaluación a los destinatarios
- recopilar datos
- comparar los datos con los objetivos del programa

b) *Modelos de enjuiciamiento que acentúan criterios intrínsecos:*

- el evaluador tiene un papel muy importante
- es un planteamiento técnico-profesional
- carece de conexión con el contexto

c) *Modelos de enjuiciamiento que acentúan criterios extrínsecos:*

— *Modelo de Scriven:*

- distinción evaluación formativa-sumativa
- se demanda especialmente el interés de los objetivos

— *Modelos de Stake* (evaluación respondiente)

d) *Modelos de facilitación de decisiones:*

— *Modelo de Cronbach*

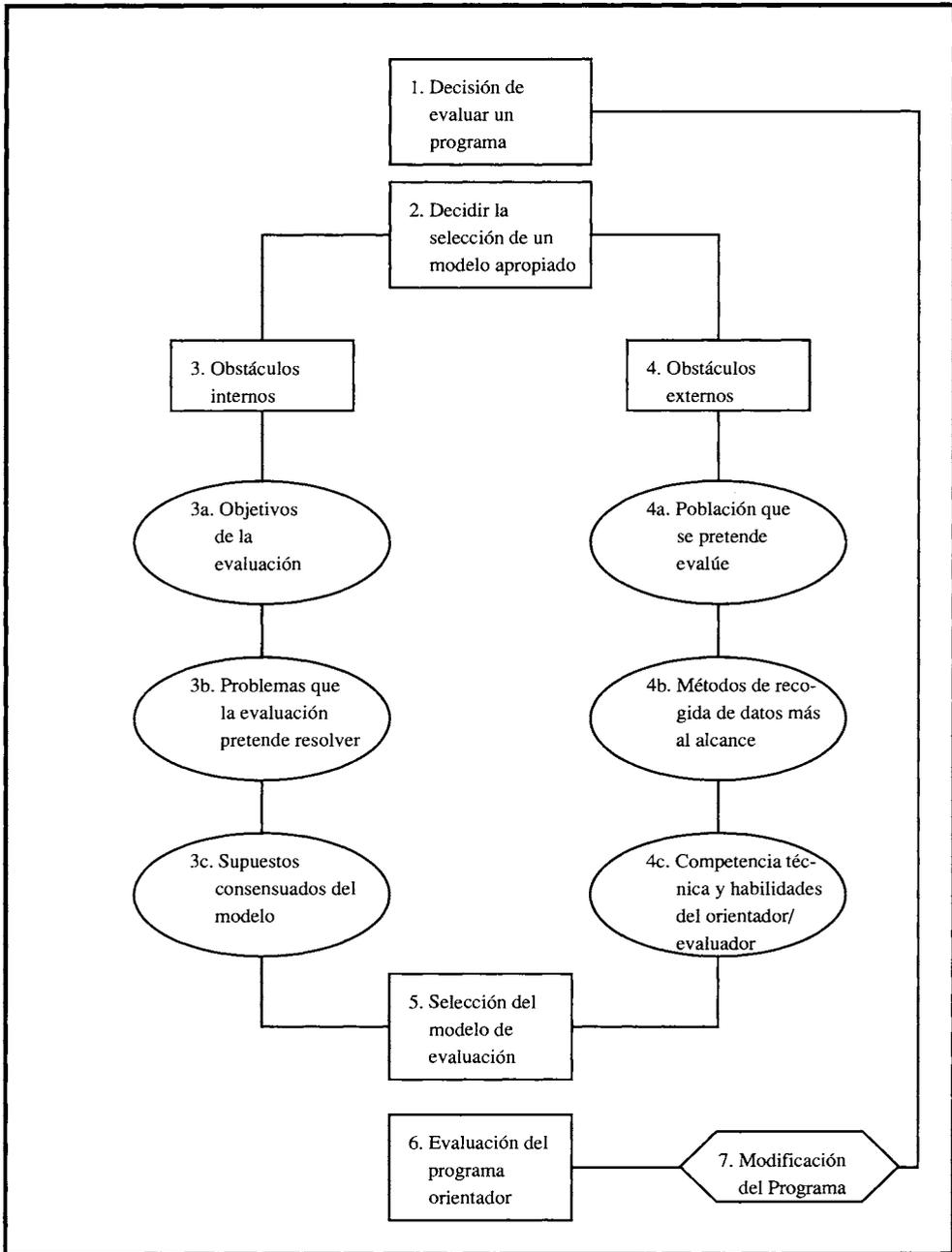
— *Modelo CIPP de Stufflebeam*

— *Modelo CSE de Alkin (variante del CIPP)*

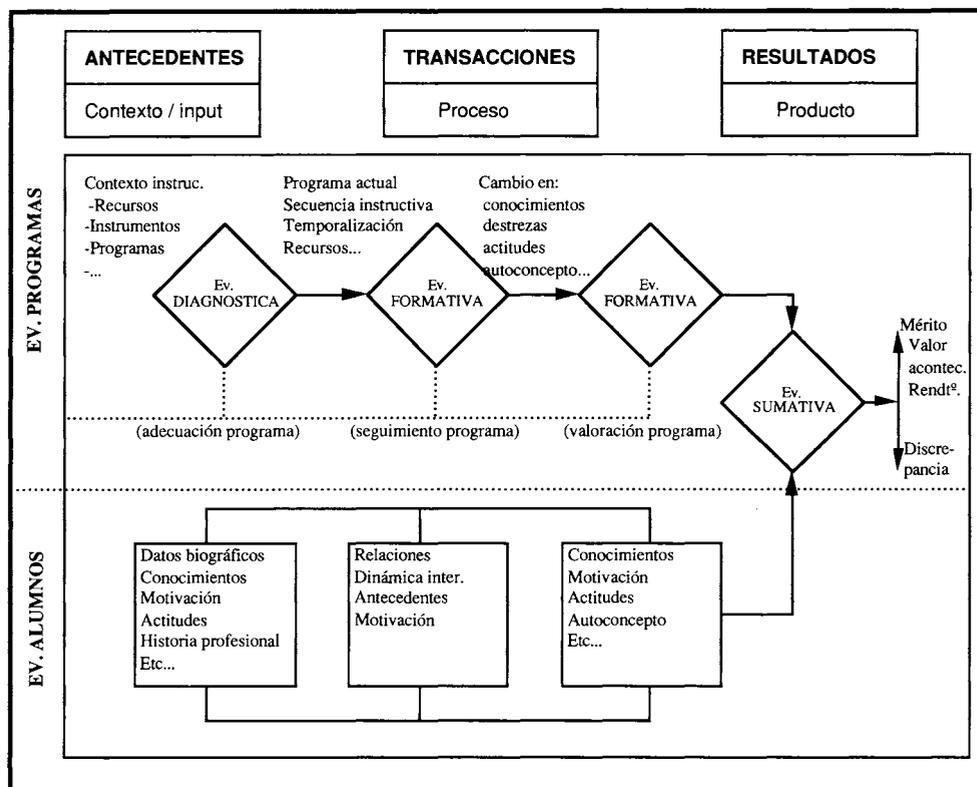
e) *Evaluación como negociación:*

— *Modelo respondiente constructivista de Guba y Lincoln*

CUADRO 3
METAMODELO DE EVALUACIÓN DE PROGRAMAS DE DANIELS (1981)



CUADRO 4
PROPUESTA DE MODELO DE EVALUACIÓN DE PROGRAMAS



La propuesta más reciente de Rebollo (1993) plantea la clasificación de los modelos en tres grandes bloques:

a) *Modelos objetivistas*, aquellos que se conciben desde una perspectiva técnica, científica. Algunos de los modelos que pueden incluirse aquí serían:

- Evaluación basada en objetivos (Tyler)
- Planificación evaluativa (Cronbach)
- C.I.P.P. (Stufflebeam)
- Evaluación sin referencia a objetivos (Scriven)

b) *Modelos subjetivistas*, que plantean la evaluación como estrategia de comprensión y valoración de los procesos y resultados de un programa. Se diferencia de la postura anterior, sobre todo, en la concepción de la realidad y en la forma de entender el conocimiento. Especialmente interesa evaluar los procesos. Podríamos incluir los modelos siguientes:

- Evaluación respondiente (Stake)
- Evaluación iluminativa (Parlett y Hamilton)
- Evaluación democrática (McDonald)

c) *Modelos críticos*, que entienden la evaluación de programas como un proceso de recogida de información que fomenta la reflexión y que persigue la transformación de los destinatarios del programa. La evaluación se centra en el análisis de las circunstancias personales, sociales y políticas de los destinatarios. El contenido básico será tanto el proceso como las actividades de su desarrollo. No hay modelos explícitamente vinculados a esta modalidad aunque si hay experiencias interesantes que siguen las pautas señaladas (Martínez y otros, 1990; Pozo y otros, 1990).

Aunque la selección de un modelo adecuado de evaluación debe realizarse en función de las características específicas del programa a desarrollar, consideramos pertinente ofrecer las pautas de carácter general que pueden ayudarnos a realizar la selección, presentadas por Daniels y otros (1981), y que configuran un metamodelo (cuadro 3).

Una propuesta concreta de modelo de evaluación, aplicada en este caso a la formación ocupacional, es la realizada por Tejada (1991) y en la que se excluye la evaluación previa a la planificación y se integran elementos teóricos de diversos modelos (Scriven, Stake y Stufflebeam). Dejamos al lector que realice su propia interpretación (cuadro 4).

5. FORMULACIÓN DE OBJETIVOS E INDICADORES

Tal y como viene planteándose la evaluación de programas, en cualquiera de sus modalidades o enfoques, la formulación de objetivos es una tarea fundamental que en buena medida va a condicionar el desarrollo del programa y las posibilidades reales de una evaluación correcta y fructífera. Por ello conviene sistematizar al máximo dicha tarea, lo que puede lograrse siguiendo cualquiera de las taxonomías o modelos de formulación de objetivos conocidos. En cualquier caso, la adecuación de estos objetivos a programas de intervención en el ámbito psicopedagógico, no exclusivamente de carácter instructivo, es una importante tarea a realizar ya que la mayoría de las taxonomías conocidas están pensadas específicamente para los procesos de enseñanza-aprendizaje.

Quizá por esta razón sigue proponiéndose (Colás, 1993) el modelo de Hammond como adecuado para la formulación de objetivos. El modelo, que aparece comentado en la referencia dada, incorpora cuatro dimensiones: programas, objetivos, acciones y agentes.

Realizada la selección de objetivos a partir de algunas pautas como las anteriormente reseñadas y antes de llegar a una propuesta definitiva, tendremos que asegurarnos de que:

- son relevantes, desde la perspectiva educativa, social, política,...., en la que se orienta el programa

- son útiles, para los sujetos a los que va destinado
- son susceptibles de ser conseguidos con la realización del programa, lo que implica que se han formulado tomando en cuenta el contexto y los recursos
- son susceptibles de ser evaluados.

Tarea nada fácil será convertir los objetivos seleccionados en indicadores. En evaluación de programas los indicadores se sitúan en la confluencia entre lo teórico y lo empírico y mientras que la definición conceptual establece el significado en términos abstractos, los indicadores se refieren a sus características observables, medibles y empíricamente detectables (como eslabón intermedio de esta cadena actuarán los objetivos operativos).

6. LOS DISEÑOS EN LA EVALUACIÓN DE PROGRAMAS

Un diseño es un plan que fija qué medidas y cuándo serán tomadas en el curso de un proceso de evaluación. Plantearse la especificación de un diseño supone haber realizado reflexiones previas en torno a cuestiones relacionadas con el concepto general de evaluación y con las finalidades de las diferentes opciones metodológicas; posteriormente delimitaremos, como parte de la planificación prevista, los fundamentos de la misma, los objetivos, los elementos personales implicados en el programa, los ámbitos de intervención, los instrumentos de recogida de datos, las técnicas de análisis y los momentos de la evaluación; todo ello, de acuerdo al modelo que se haya elegido para realizar la actividad evaluativa y teniendo en cuenta que en todos estos elementos existen diferencias según la perspectiva o enfoque (paradigma o modelo) por el que se haya optado.

En cualquier caso, independientemente del modelo que se haya elegido, cualquier mínima planificación tiene que atender las siguientes tareas:

- establecer las actividades que se van a evaluar
- fijar los criterios de evaluación
- elegir las estrategias para la obtención de información
- analizar la información
- tomar decisiones a partir de la información analizada

En su concreción, la variedad de diseños es amplísima ya que no hay limitación alguna en la elección, si bien dadas las características peculiares de la investigación evaluativa existen diseños más adecuados para planificar la intervención mediante programas. La propuesta de Fitz-Gibbon y Morris (1987, pp. 53-60) hace referencia a siete modalidades de diseño:

- 1) Diseño pretest post-test con grupo de control
- 2) Diseño post-test (únicamente) con grupo de control
- 3) Diseño pretest post-test con grupo de control no equivalente
- 4) Diseño de series temporales con un único grupo (experimental solamente)
- 5) Diseño de series temporales con grupo de control no equivalente
- 6) Diseño antes-después

7) Diseños complejos (experimentales y cuasiexperimentales, resueltos mediante análisis de varianza).

Creemos que esta propuesta olvida la referencia a determinadas estrategias descriptivas y correlacionales, razón por la cual sugerimos la siguiente:

A) *Diseños para la investigación cuantitativa*

- 1) Diseños descriptivos-comparativos (estudios tipo survey)
- 2) Diseños longitudinales
 - secuenciales (diacrónicos)
 - transversales (sincrónicos)
- 3) Diseños correlacionales:
 - correlación y regresión (discontinuidad en la regresión)
 - técnicas de clasificación
- 4) Diseños causales
- 5) Diseños experimentales:
 - a) preexperimentales
 - pretest con un solo grupo
 - pretest post-test con un solo grupo
 - b) cuasiexperimentales
 - pretest posttest con grupo control no equivalente
 - series con pretest y posttest con un solo grupo
 - series con grupo control
 - c) experimentales de campo
 - pretest con grupo control
 - pretest posttest con grupo control
 - n=1 (línea base)
 - intrasujetos
 - complejos

B) *Diseños para la investigación cualitativa*

- 1) Estudios de casos
- 2) Estudios del grupo
- 2) Estudios observacionales
- 3) Estudios de comunidades (etnográficos)
- 4) Estudios de situaciones o contextos
- 5) Estudios de investigación-acción

Nos parece importante insistir en esta distinción ya que tradicionalmente el diseño de evaluación se ha asociado a la práctica de evaluaciones sumativas, de carácter cuantitativo. Mientras que la evaluación formativa no suele planificarse de acuerdo a pautas de diseños cerrados, lo que no quiere decir que este tipo de

evaluación no requiera reflexiones en torno a la organización del trabajo a desarrollar.

Hay campos de aplicación de la acción educativa (educación compensatoria, educación especial) en los que la planificación de diseños supone muchas más dificultades que en otros (educación de adultos, proceso instructivo,...). En cualquiera de los campos, es fundamental admitir la diversidad de aplicaciones en función de la modalidad de evaluación: cabe pensar en un diseño para la evaluación de necesidades, en otro para la evaluación de contextos y en otro para la evaluación de resultados. Igualmente será importante el campo específico de trabajo, con ciertos tipos de diseños más adecuados para programas de educación especial, por ejemplo, y diseños más propios para programas de procesos instructivos.

7. FASES DEL PROCESO EVALUATIVO

Presentamos las fases correspondientes al diseño de programas de corte tradicional y de corte participativo.

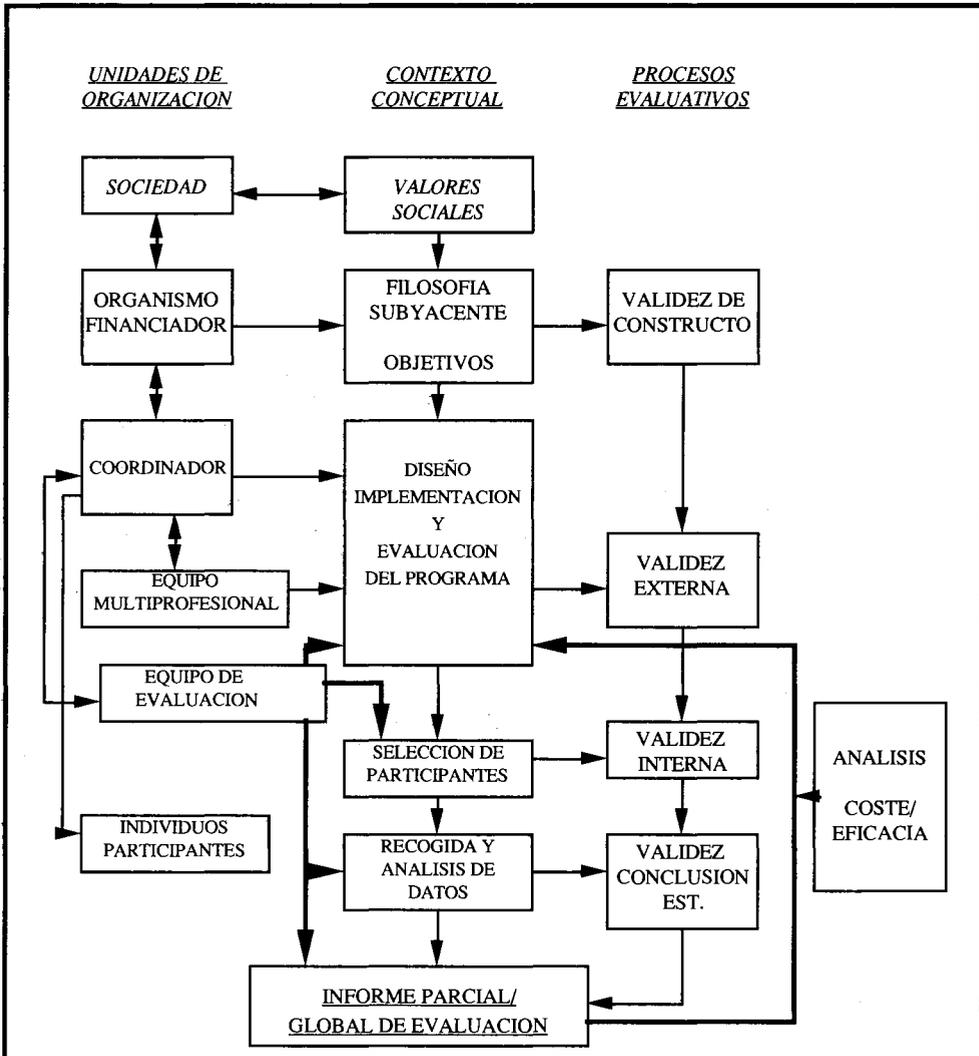
A) *Esquema tradicional*

El esquema de planificación de un programa incluiría las siguientes fases:

- 1) Análisis o evaluación de necesidades
- 2) Programación:
 - desarrollo de objetivos y procedimientos de evaluación
 - selección de estrategias o actividades para lograr los objetivos
- 3) Implementación:
 - desarrollo de estrategias
 - construcción de instrumentos, que implica:
 - delimitación de variables e indicadores
 - selección y listado de criterios
 - redacción de ítems
 - aplicación de instrumentos
 - seguimiento del proceso
- 4) Toma de decisiones (evaluación propiamente dicha):
 - sintetizar, analizar y valorar la información
 - proceso dinámico que genera información
 - retroalimentación del programa
 - carácter formativo vs sumativo

Un esquema de corte tradicional similar al presentado es el ofrecido por Alvira (1991, p. 22). Algo más exhaustivo y orientador de las actividades a llevar a cabo en la puesta en marcha de un programa es el modelo de Wortman (1983), que presentamos a continuación (cuadro 5).

CUADRO 5
MODELO DE WORTMAN



B) *Esquema participativo*

Las dos peculiaridades principales que desde un principio incorpora la investigación evaluativa respecto a otros esquemas más tradicionales de investigación (básica, aplicada,...), son, como ya hemos señalado, la participación de los interesados en algunas de las fases del proceso de evaluación y la utilización de los resultados para

la mejora del propio programa (feed-back para la toma de decisiones). Sin embargo, no siempre la investigación evaluativa ha respondido a estas expectativas, por lo que desaparecen los matices diferenciales de esta estrategia de investigación en relación con otras de corte más clásico.

Potenciar explícitamente esos matices como criterios de identificación más específicos de la investigación evaluativa es lo que pretenden conseguir algunas de las últimas propuestas realizadas (el «modelo de evaluación responsable» de Stake y el «modelo de evaluación dirigida a resultados» de Patton).

La investigación evaluativa tiene que preocuparse de proporcionar a los interesados la información que necesitan, estudiando los problemas que ellos consideran relevantes. El hincapié en estos aspectos no altera el flujo (proceso/fases) de una evaluación, pero sí cambia el énfasis y contenido de cada fase, de modo que la identificación de los problemas (objetivos) a través de contactos con los interesados (audiencia) y la consideración de los distintos métodos para establecer el consenso entre las partes, toma preminencia sobre otras fases de la investigación.

A partir de estos considerandos, y tomando como referencia el modelo de Patton, en el trabajo ya comentado de Alvira (1991, p. 31), nos presenta el esquema de actuación para el desarrollo de una investigación evaluativa que incorpora explícitamente los presupuestos participativos y que por su interés presentamos aquí al lector sin comentarios añadidos.

- 1) Identificación de partes interesadas
- 2) Organización con ellas de un grupo de trabajo
- 3) Delimitación de preguntas y temas de la evaluación
- 4) Anticipación de las respuestas y valoración de las mismas a fin de determinar si procede seguir con la evaluación
- 5) Decisiones sobre diseño, métodos y medición
- 6) Recogida de información
- 7) Análisis de datos
- 8) Implicación en el análisis del grupo de trabajo
- 9) Decisión sobre cómo utilizar y aplicar los hallazgos
- 10) Decisión sobre difusión de los hallazgos

8. TIPOS DE EVALUACIÓN EN FUNCIÓN DE LAS ETAPAS DEL PROGRAMA

8.1. Etapas en la Evaluación de Programas

Pueden establecerse básicamente tres etapas: evaluación de la formulación del programa; evaluación de la aplicación del programa y evaluación de los resultados del programa.

Estas diferentes etapas de la evaluación pueden presentarse como diferentes tipos de evaluaciones. Angulo (1988), como ya expusimos en un trabajo anterior (Tejedor, 1990, p. 20), relaciona las distintas inquietudes planteadas en el desarrollo

CUADRO 6
ESTRUCTURA METODOLÓGICA

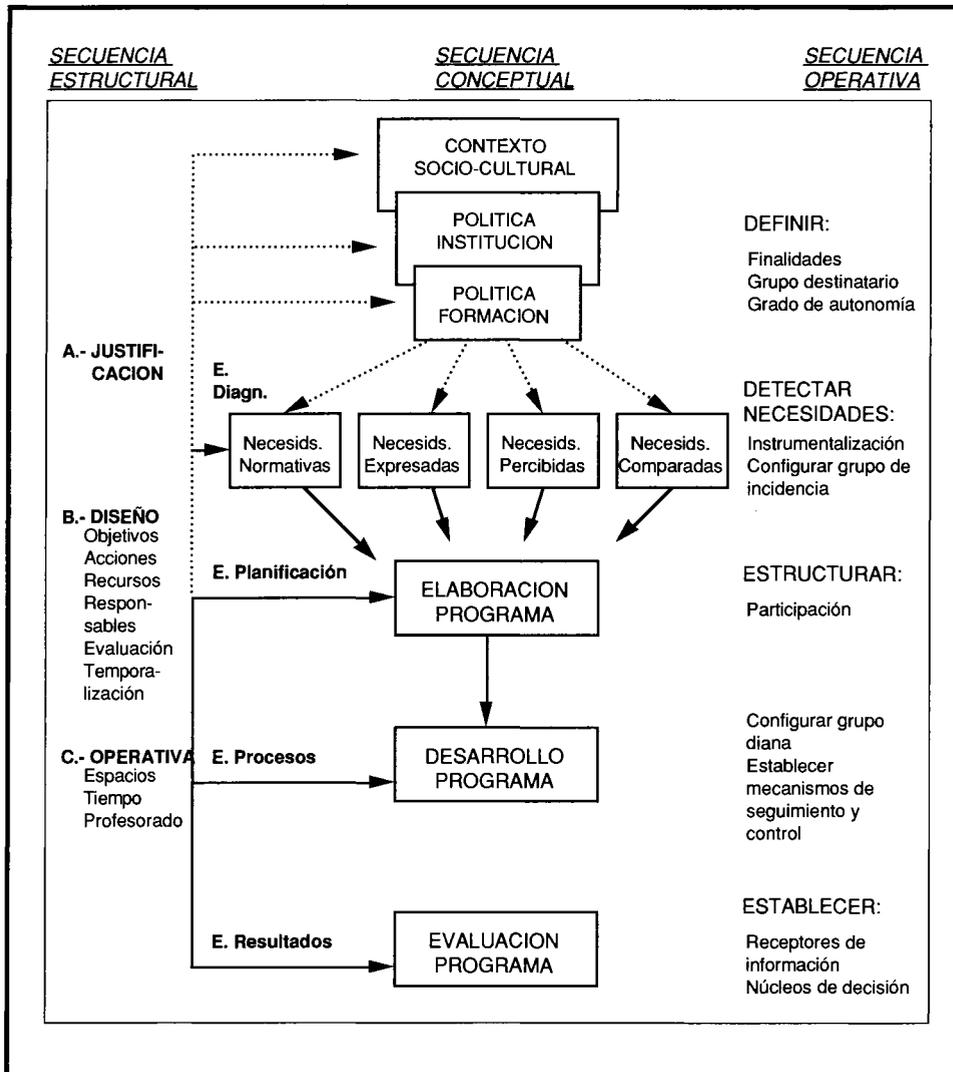
Estructura metodológica de la evaluación en el desarrollo comunitario			
Tipo de evaluación	Referente	Tarea	Ambito de aplicación
Del contexto necesidades	territorio y sociedad: comunidad, grupos, sujetos,...	estudio previo de la realidad:problemas, necesidades, recursos, conflictos,...	diagnóstico, planificación
Del diseño	expertos y profesionales en programar y planificar	análisis de criterios programáticos, justificación coherencia y pertinencia del diseño	diseño y planificación del programa
Del proceso	profesionales participantes en programa	análisis desarrollo del programa: estrategias, niveles de ejecución, participación,...	ejecución de programa
Del producto	equipo evaluador	valoración de los resultados en relación con objetivos:eficacia, eficiencia,...	conclusiones toma de decisiones

del programa con los diferentes tipos de evaluaciones. Podemos pensar también que, como recogemos en el cuadro 6, cada etapa del programa lleva asociada un tipo de evaluación (Caride, 1990), con especificación de ámbitos, tareas y referentes de evaluación.

Recientemente Alkin (1991) identificaba cinco tipos de evaluación vinculados a las distintas fases en la evaluación de programas:

- 1) Evaluación de sistemas (fase de preplanificación o de evaluación de necesidades)
- 2) Evaluación de planificación del programa (evaluación del diseño)
- 3) Evaluación de la puesta en marcha del programa (evaluación de proceso)

CUADRO 7
SECUENCIA DE ACTIVIDADES EN EVALUACIÓN DE PROGRAMAS



- 4) Evaluación de perfeccionamiento del programa (evaluación de resultados)
- 5) Evaluación de certificación del programa (evaluación de eficacia y evaluación de impactos).

Una panorámica óptima de la secuencia de actividades a desarrollar para la elaboración de programas es la que nos presenta Gairín (1991), que nosotros recogemos en el cuadro 7.

8.2. Contenidos o referentes de la evaluación

Desde una perspectiva metodológica, la puesta en marcha de un programa supone una actuación que debe atender ordenadamente a las distintas fases del proceso. La supervisión del desarrollo de cada una de esas fases en términos evaluativos nos lleva al establecimiento de diferentes tipos de evaluación que pueden llegar a tener interés en sí mismos, incluso con su metodología propia (necesidades, objetivos, proceso, resultados,...); pero no podemos olvidar que todos ellos forman parte de un esquema único de investigación que nos informará sobre la puesta en marcha del programa y la calidad de sus logros.

Tenemos entonces, de acuerdo con la fase del proceso a la que se atiende, la posibilidad de establecer diversos tipos de evaluación:

1) *Evaluación de necesidades*

Debe ser la primera actividad a llevar a cabo en la puesta en marcha de un programa de intervención. Aunque, como ya señalamos en escritos anteriores (Tejedor, 1990), algunos autores plantean la evaluación de necesidades con el mismo énfasis que la evaluación de programas, considero que es más pertinente concebirla como una fase de dicho proceso, limitando entonces sus objetivos, lo que no implica desvirtuar su esquema operativo que entiendo está suficientemente conformado.

Es decir, podemos pensar, desde un punto de vista metodológico, que la evaluación de necesidades, siendo una fase del proceso de investigación para la puesta en marcha de un programa de intervención, tiene entidad propia y es susceptible de ser analizada con un esquema completo de investigación evaluativa (Bank y Morris 1979; Witkin, 1984).

2) *Evaluación del diseño del programa de intervención*

Supone la preocupación por la secuencia de los pasos a seguir, por la claridad de formulación de los objetivos planteados, por la calidad de los instrumentos o técnicas de obtención de información, por la suficiencia de la muestra fijada, por la adecuada localización de las unidades muestrales, por la procedencia de las actividades sugeridas, por la adecuación de las técnicas de análisis,... (Mayer, 1985; Bartolomé, 1990).

3) *Evaluación de la evaluabilidad*

La experiencia determina que pocos de los programas llevados a la práctica son evaluables al tener problemas de especificación de objetivos, de implementación, de adecuación de las actividades a la consecución de objetivos, de carencia de medios para su desarrollo,... Por ello estamos de acuerdo con quienes sugieren la conveniencia de llevar a cabo una revisión previa de las posibilidades evaluativas del

CUADRO 8
EVALUABILIDAD DE MODELOS

El decisor político asigna recursos para la evaluación de un programa
Determinación de la evaluabilidad (actividades)
<p>Determinación de la evaluabilidad (productos)</p> <ol style="list-style-type: none"> 1) Objetivos y expectativas 2) Medidas del desempeño 3) Factibilidad de medir el desempeño 4) Usos probables de la información 5) Opciones de evaluación
<p>Uso de la información de evaluación</p> <ol style="list-style-type: none"> 1) Mantener o cambiar los objetivos 2) Mantener o cambiar las actividades 3) Mantener o cambiar las estrategias de desarrollo del programa 4) Seleccionar una o más de las opciones de evaluación

programa, lo que supone fundamentar adecuadamente las pautas metodológicas a seguir, dado que, al igual que la evaluación de necesidades, es un campo de referencia con tratamiento específico ya contrastado (Rutman, 1980; Shipman, 1989; Anguera, 1990; Berk y Rossi, 1990).

Wholey (1979) presenta un esquema para determinar la evaluabilidad de un programa (cuadro 8).

4) Evaluación de la implementación

Este tipo de evaluación pretende analizar el funcionamiento del programa y determinar si existen diferencias entre el funcionamiento real y el previsto inicialmente. Si hubiera diferencias notorias entre ambos esquemas de funcionamiento

trataríamos de hallar las causas y, si fuera posible, subsanar las deficiencias (Leithwood y Montgomery, 1980; Finnegan y otros, 1988).

5) *Evaluación de la cobertura*

Trataremos de analizar si el programa cubre todos los objetivos previstos y si llega a todos los sujetos para los que se había pensado, determinando, si procede, los sesgos de cobertura existentes y las causas que los han motivado.

6) *Evaluación del proceso (seguimiento)*

En algunos casos, sobre todo si el programa es de larga duración, conviene realizar un análisis de su funcionamiento a fin de verificar si su desarrollo sigue las pautas previstas. Implica la revisión de todos los elementos constitutivos del programa (personales, funcionales y materiales), en su dimensión interactiva: calidad de las actividades, distribución del tiempo y del trabajo, utilización de recursos,...

Pretende describir y valorar las actividades que se llevan a cabo durante el desarrollo del programa intentando solucionar los fallos que pudieran encontrarse.

La evaluación adecuada del proceso, en tanto en cuanto que implica que el programa se ha puesto en marcha, supone la consideración de todos sus elementos por lo que se sugiere un esfuerzo previo de reconsideración respecto a las fases anteriores de ejecución del programa (necesidades, objetivos, diseño,...).

7) *Evaluación de resultados*

Hasta hace relativamente poco tiempo era el único tipo de evaluación que se llevaba a cabo tras la finalización de las actividades que implicaba la ejecución del programa.

En términos generales, se trata de analizar el grado en que el programa consigue los resultados previstos (también pueden evaluarse, siempre que se hayan previamente identificado, resultados no previstos inicialmente). Se pretende determinar hasta qué punto pueden vincularse unívocamente los logros obtenidos con los efectos del programa, lo que nos sitúa de lleno en la problemática de la validez interna (control de variables) de todo trabajo de investigación, cuestión que, como es sabido, tiene especial interés en los diseños de intervención dada la dificultad para establecer nexos de causalidad y covariación entre las variables consideradas en el programa.

La evaluación de resultados admite la modalidad de «evaluación de verificación», que pretende efectuar comprobaciones a base de patrones claramente identificados, tanto en base a datos cuantitativos como cualitativos, y la modalidad de «evaluación comprensiva», que pretende ampliar o profundizar la comprensión de situaciones y que utiliza básicamente datos cualitativos.

La temática de la evaluación de resultados es amplia y compleja, apareciendo tratada en la bibliografía con diferentes terminologías: evaluación de productos, de la efectividad, de la eficacia, de la eficiencia, ... Las propuestas metodológicas (conceptualización del producto, selección de productos relevantes, diseños, instrumentos, medida y análisis de datos) para el análisis de resultados son igualmente muy variadas (Judd y Kenny, 1981; De la Orden, 1985 a, 1990).

En algunos casos, la evaluación de resultados se plantea en términos de «evaluación del impacto», aunque la mayoría de los estudiosos del tema reserven esa terminología para referirse a la evaluación de resultados en la población del entorno de los usuarios del programa; es decir, se trataría tanto de analizar las consecuencias no previstas al planificar el programa como de evaluar su incidencia en una población más amplia que la directamente receptiva del programa.

La estimación del impacto nos permite identificar y prevenir las consecuencias de un programa y desemboca en la valoración de los efectos, esperados o no. Las técnicas más utilizadas en la estimación de impactos son, en opinión de Anguera (1989):

- las matrices de revisión causa-efecto, para la búsqueda de relaciones entre las variables (para cada una de las actividades básicas de intervención se revisan los factores ambientales y/o sociales que puedan verse afectados)
- matrices cruzadas, para el establecimiento de relaciones primarias, secundarias,...
- modelos de simulación
- análisis de supervivencia
- estimación de demandas futuras
- modelos logit de respuestas multinómicas

8) *Evaluación económica*

Trata de analizar el coste del programa, tanto en términos absolutos como en términos relativos (respecto a resultados conseguidos).

El análisis de los costes surge en el ámbito de los estudios de las ciencias económicas y se plantea en relación directa con los beneficios. Como en la intervención social no tiene mucho sentido hablar de «beneficios», el estudio de los costes se plantea en términos de análisis «coste/efectividad» o «coste/utilidad», lo que vincula este tipo de evaluación a la evaluación de resultados, debiendo entenderse como una consecuencia de ella y, por tanto, realizarse con posterioridad (Levin, 1985).

9. LOS CRITERIOS Y NORMAS DE VALOR

Sin duda, uno de los problemas básicos que rodean a toda actividad evaluativa es el problema del establecimiento de los criterios, que nos van a permitir la emisión de juicios de valor respecto a la consecución o no de los objetivos planteados a

partir de las actividades llevadas a cabo. Es un proceso en el que pueden señalarse tres fases:

- 1) Especificación de los criterios que nos señalen lo que tiene que hacer o cómo tiene que funcionar un programa para ser etiquetado como bueno.
- 2) Para cada criterio hay que especificar estándares de funcionamiento que detallen niveles o grados de mérito.
- 3) La evaluación consistirá simplemente en la recogida de información respecto a esos criterios para estimar si se han alcanzado o no los estándares prefijados.

No siempre es fácil establecer criterios y estándares pero su determinación es un requisito previo para una adecuada evaluación. Desde luego, la formulación clara de objetivos, y más si éstos están debidamente operacionalizados, facilitará la tarea del establecimiento de los criterios (en algún caso, los objetivos así establecidos se identifican con los propios criterios).

Brunet (1987) señala los cuatro aspectos principales que hay que controlar para lograr que la medición del criterio sea lo más válida posible:

- la insuficiencia del criterio (comprobar si el criterio que se quiere medir incluye todos los objetivos deseados del programa)
- la contaminación del criterio (se refiere al control de ciertos elementos externos del programa que pueden afectar a los resultados evaluados)
- los sesgos causados por las características del grupo, que se producen cuando los participantes en el programa poseen ya las características medidas por el criterio
- la pluridimensionalidad del criterio, lo que exige evaluar el criterio con múltiples medidas.

Van de Ven y Ferry (1980) proponen una guía de seis puntos para llevar a cabo la tarea de fijación de los criterios de evaluación:

- 1) Deben escogerse criterios que tengan gradaciones para que puedan ser discriminativos.
- 2) Como es imposible medir todos los criterios que podemos considerar como relevantes de las distintas variables, conviene investigar aquellos que parecen explicar o están relacionados con el mayor número de variables (la parsimonia es por tanto un aspecto a tener en cuenta).
- 3) Se han de buscar aquellos criterios que son menos costosos de medir, tanto económicamente como en el tiempo.
- 4) Hay que determinar si cada criterio se asocia con una variable dependiente, independiente o moduladora.
- 5) Hay que clasificar los criterios por niveles de análisis.
- 6) Tanto los usuarios como los evaluadores han de realizar una exhaustiva revisión de los criterios obtenidos hasta alcanzar el consenso sobre los más adecuados.

Aunque cada programa deberá formular sus propios criterios, podemos hablar de criterios de valor de referencia general: utilidad, viabilidad, formalidad, exactitud

CUADRO 9
CRITERIOS E INDICADORES DE EVALUACIÓN

<u>CRITERIOS</u>	<u>INDICADORES</u>
<u>Pertinencia:</u> Adecuación de un programa con la política de formación y el contexto de aplicación.	<i>Nivel de coherencia: relación entre los objetivos asignados y los objetivos propuestos.</i>
<u>Actualización:</u> Adecuación de los objetivos del programa y las necesidades reales (sociales e individuales).	<i>Relación entre los objetivos propuestos y las necesidades detectadas (objetivos de formación-demanda/oferta laboral).</i>
<u>Objetividad:</u> Adecuación de las leyes y principios científicos.	<i>Relación entre los objetivos asignados y contenidos (selección y secuenciación de contenidos).</i>
<u>Aplicabilidad:</u> Posibilidad de puesta en práctica de los objetivos propuestos.	<i>Relación entre el programa y la inserción laboral.</i>
<u>Suficiencia:</u> Grado con que un programa satisface las necesidades detectadas.	<i>Nivel de exhaustividad: relación entre los objetivos asignados y las necesidades detectadas.</i>
<u>Eficacia:</u> Nivel de logro de los objetivos asignados.	<i>Relación entre los objetivos asignados y los objetivos alcanzados.</i>
<u>Eficiencia:</u> Grado de implicación de recursos humanos, materiales y funcionales.	<i>Relación entre los objetivos logrados y los recursos implicados (ratio profesor/alumno, hora/profesor/alumno, etc.)</i>
<u>Comprensividad:</u> Grado de optimización alcanzado.	<i>Relación entre el nivel de entrada y nivel de salida de un programa.</i>
<u>Relevancia:</u> Grado de importancia del programa para cubrir necesidades individuales y sociales.	<i>Relación entre los objetivos propuestos y necesidades sociales e individuales (objetivos de formación-necesidades, expectativas, intereses, cambio laboral y polivalencia laboral, etc.)</i>

(precisión y validez), contextualización y diversidad en la valoración de las consecuencias/efectos del programa, juzgadas desde el punto de vista de los técnicos, de los gestores y de los usuarios.

Ferrández (1990) realiza una interesante propuesta en la que aparecen detallados criterios e indicadores, con un matiz operativo y referido a la puesta en marcha de un programa de formación ocupacional (cuadro 9).

Desde una perspectiva más global (considerando la evaluación del programa en su totalidad), también podemos hablar, de forma genérica, de criterios o normas para la realización de un proceso evaluativo.

Las normas que deben orientar las realizaciones de las prácticas evaluativas, definidas como principios comúnmente aceptados para determinar el valor o la calidad de una evaluación, han sido fijadas por colectivos de investigadores, pudiendo citarse como referencias claves los acuerdos del Joint Committee on Standards for Educational Evaluation de 1975 y de 1981 (Standards for Evaluation of Educational Programs, Projects and Materials, publicadas bajo la supervisión de Stufflebeam).

Su establecimiento pretende proporcionar, en opinión del propio Stufflebeam (1991):

- definiciones operativas para la evaluación de programas y de personal
- grados de acuerdo respecto a métodos y significados de la evaluación educativa
- estrategias para afrontar los problemas evaluativos
- guías prácticas para planificar las evaluaciones
- criterios aceptados para valorar los planes de las evaluaciones
- referentes conceptuales para fundamentar la evaluación
- pautas para llegar a la profesionalización de la tarea evaluadora
- diversos puntos de vista sobre los diferentes tipos de evaluación
- sugerencias para formar a los evaluadores.

Una línea similar siguen las conocidas 95 tesis para la reforma de la práctica evaluativa de Cronbach (1980).

10. TÉCNICAS DE RECOGIDA Y ANÁLISIS DE LA INFORMACIÓN

10.1. Técnicas de recogida de datos

La información que se precisa en cada tipo de estudio evaluativo es muy diversa y por tanto el único consejo válido al evaluador es seleccionar adecuadamente de entre el amplio espectro de técnicas las que sean más adecuadas en función de los objetivos del estudio y de las posibilidades reales del mismo, recordando que no hay técnicas específicas aunque sí existen algunas más adecuadas que otras para recoger información pertinente según la modalidad de evaluación que se está llevando a cabo.

Sin pretender ser exhaustivos, recogemos en el cuadro 10 los instrumentos adecuados en función del aspecto a evaluar.

CUADRO 10
INSTRUMENTOS DE MEDICIÓN

<p>1) Características de los sujetos</p> <p>inteligencia-aptitudes</p> <p>personalidad</p> <p>rendimiento</p> <p>actitudes, opiniones, intereses</p> <p>datos sociológicos</p>	<p>test estandarizados</p> <p>test y cuestionarios estandarizados</p> <p>pruebas estandarizadas y no estandarizadas</p> <p>escalas, cuestionarios, entrevistas, encuestas diarias</p> <p>entrevistas, cuestionarios</p>
<p>2) Características sociales del grupo</p> <p>relaciones del grupo</p> <p>contexto económico y social</p> <p>sistemas organizativos</p>	<p>test sociométrico</p> <p>observación</p> <p>incidente crítico</p> <p>escalas BARS</p> <p>indicadores, cuestionario</p> <p>documentos, encuesta</p> <p>entrevistas</p> <p>técnica delphi</p> <p>técnicas negociación</p>

No pensamos que la utilización de estas técnicas de obtención de información en el campo de la evaluación de programas tenga que responder a criterios diferentes a los establecidos para su utilización en otros campos de la investigación social. Sin duda, lo más importante a tener en cuenta desde un principio será adaptar el proceso de construcción de instrumentos (cuestionarios, entrevistas, escalas de observación,...) al tipo de evaluación que se quiera llevar a cabo (descriptiva, comprensiva); a los objetivos y condiciones específicas del trabajo a realizar; conocer a quién van

dirigidos (padres, alumnos, profesores, autoridades,...) y las posibilidades de aplicación (amplitud de la muestra, disponibilidades económicas,...); y disponer de los recursos necesarios para codificar y analizar adecuadamente la información recogida así como para difundir los informes pertinentes.

En escritos anteriores hemos comentado las peculiaridades de estas técnicas en su aplicación a la evaluación de necesidades (Tejedor, 1990). Pensamos que aquellas opiniones son extensibles al marco más general de la evaluación de programas por lo remitimos al lector interesado a dicho escrito.

10.2. Técnicas de análisis de datos

Insistiendo en algunas opiniones ya expuestas en escritos anteriores (Tejedor, 1990), creemos que con las técnicas a utilizar en el análisis de datos de la evaluación de programas ocurre otro tanto que lo señalado en relación a los instrumentos; no existen técnicas específicas pero conviene seleccionar las que mejor se ajusten a los presupuestos metodológicos implicados: tipo de evaluación seleccionada, peculiaridades del diseño, instrumentos utilizados en la recolección de datos, características de las variables y tipo de información esperada.

Las consideraciones a tener en cuenta para la elección de la técnica concreta de análisis incluirá referencias a:

- adecuación de la técnica al enfoque dado en la investigación
- posibilidad de evaluar los objetivos formulados a partir de los datos procedentes de los análisis sugeridos
- adecuación al tipo de datos obtenidos (por tanto, posibilidad de utilizar distintas técnicas si tenemos distintos tipos de datos)
- disponibilidad de la infraestructura para el tratamiento adecuado de la información (análisis de entrevistas, número de observadores, soporte informático,...).

Es un campo en el que no se puede, a priori, pensar en la improcedencia de aplicación de técnica alguna, ya que se mueve entre la aplicación del análisis de diarios y la de los modelos matemáticos rigurosos; que puede utilizar desde las técnicas de observación menos estructuradas a los diseños experimentales de campo; que puede incorporar datos de muy variada naturaleza,... Stufflebeam y otros (1984, pp. 88-89) presentan una relación interesante de las técnicas que pueden asociarse con este tipo de estudios.

Aún aceptando este contexto de diversidad, pienso que pueden señalarse algunas técnicas como más específicamente adaptables al análisis de datos en el estudio de la evaluación de programas:

- a) Todas aquellas que permiten comparar resultados o situaciones diversas, bien por la vía del establecimiento de diferencias significativas bien por la vía del reconocimiento de ganancias o pérdidas.
- b) Aquellas técnicas adaptables al análisis de contingencia, tanto bidimensiona-

les como multidimensionales. En este sentido consideramos del todo pertinente las aportaciones recientes en este campo que nos posibilitan desde el estudio de relaciones simétricas o asimétricas al estudio de modelos descriptivos (log-linear) o explicativos (logit); en todos los casos con posibilidad de estimar los parámetros del modelo, de analizar las interacciones entre las variables y de verificar el ajuste del modelo respecto a la hipótesis de trabajo planteada.

- c) Aquellas técnicas que están específicamente pensadas para el establecimiento de tipologías, tanto de variables como de sujetos (análisis clusters) y que actualmente pueden adaptarse a datos correspondientes a escalas de medida muy diversas (métricas o no métricas).
- d) Aquellas técnicas cuyo objetivo básico es el establecimiento de relaciones de covariación y causalidad. Este grupo de técnicas (regresión y análisis causal) merecen una mención específica ya que pueden permitirnos —el análisis causal— no sólo detectar las relaciones entre las variables sino establecer las causas que las originan. Para algunos investigadores, el análisis causal incluso sobrepasa el carácter de técnica específica para alcanzar el nivel de modelo, de estrategia. Interés particular puede tener el método de análisis causal denominado Fault Tree Analysis (FTA), tanto en su dimensión cualitativa como cuantitativa y que tanta utilización provechosa ha tenido en el estudio del establecimiento de causas que han orientado la toma de decisiones, la solución de problemas y la adopción de prioridades (Witkin, 1982).
- e) Las técnicas vinculadas al estudio de datos de carácter longitudinal, en sus dimensiones diacrónica y sincrónica (diseños secuenciales, time-lag, transversales, secuencial-transverso,...).
- f) Por último, podemos hacer referencia a aquellas técnicas que se vinculan de una forma específica con los procesos de toma de decisiones, es decir, con la elección de una determinada opción entre diversas alternativas. Cada opción dará lugar a un resultado distinto que puede ser medido en términos de utilidad, coste, beneficio o cualquier otra magnitud, dependiendo del problema que se esté considerando. Nos va a interesar reducir la incertidumbre que conlleva la elección de una opción, lo que se conseguirá a medida que seamos capaces de prevenir.

Desde un punto de vista metodológico, los métodos de previsión pueden agruparse en dos grandes bloques: cualitativos y cuantitativos. Los cualitativos se utilizan cuando el pasado no proporciona una información directa sobre el fenómeno considerado; las técnicas estadísticas desempeñan en ellos un papel secundario y lo más importante es contar con expertos dotados de intuición y sagacidad. Podríamos incluir entre estas técnicas el brainstorming, el delphi y el cross-impact.

En las posiciones de carácter cuantitativo se parte del supuesto de que se tiene registrada información sobre el pasado, generalmente en forma de series temporales (conjunto de observaciones acerca de una variable, observada a intervalos regulares

de tiempo). La misión del estadístico consiste en extraer toda la información posible contenida en los datos y, en base al patrón de conducta seguida en el pasado, realizar conjeturas sobre el futuro. En el análisis de series de datos de una única variable suele interesarnos conocer sus componentes (tendencias, factores cíclicos, movimientos estacionales y movimientos irregulares), de forma global o local (alissados exponenciales). Si el objetivo, más complejo, es identificar los modelos generados de las observaciones y, una vez verificado el modelo, predecir valores futuros, estaremos en el campo de los modelos ARIMA (utilización plena del análisis de series temporales).

Estas técnicas apenas si se han incorporado a la investigación educativa; quizá la consideración rigurosa del estudio de la evaluación de programas nos exija una mayor familiarización con ellas si no queremos depender excesivamente de otros investigadores que sí las han incorporado a este tipo de estudios con probada utilidad.

REFERENCIAS SOBRE PRÁCTICAS DE EVALUACIÓN DE PROGRAMAS LLEVADAS A CABO EN NUESTRO PAÍS

Recogemos, sin ánimo de ser exhaustivos, algunas referencias sobre trabajos de evaluación de programas de intervención llevados a cabo en España en los últimos años.

- ABARCA, P. (coord.) (1989): *La evaluación de Programas Educativos*. Escuela Española, Madrid. Esta publicación recoge los trabajos presentados a la reunión anual de AEDES, celebrada en Madrid, con aportaciones a la evaluación de programas en distintos ámbitos psicopedagógicos.
- AGUILERA, M^a J. y otros (1990): *Evaluación del programa de integración escolar*. CIDE, Madrid.
- ÁLVARO, M. y otros (1990): *Evaluación externa de la reforma Experimental de las E.E.M.M. (II)*: CIDE, Madrid.
- ALVIRA, F. (1989): Los centros ocupacionales y los centros de recuperación de minusválidos físicos del INSERSO. Un ejemplo de evaluación externa. *Anales de Psicología*, nº 5, 67-76.
- BARRIO, V. del (1989): Diseño de evaluación de un Programa de Educación Vial. *Anales de Psicología*, nº 5, 77-102.
- DUEÑAS, M. L. (1993): *Evaluación de un programa de apoyo a la integración escolar de deficientes mentales* (Madrid): Citado y comentado en Colás y Rebollo (1993).
- GARCÍA PASTOR, C. y GARCÍA JIMÉNEZ, E. (1993): *Evaluación de los efectos de la integración escolar en niños deficientes visuales* (Sevilla): Citado y comentado en Colás y Rebollo (1993).
- MARTÍNEZ, A. y otros (1993): *Evaluación de los programas de intervención en las instituciones de menores* (Valencia): Citado y comentado en Colás y Rebollo (1993).
- VARIOS (1989): *Evaluación de la Integración Escolar (2º Informe)*: MEC. Dirección General de Renovación Pedagógica. Madrid.
- VARIOS (1990): *Evaluación del Programa de Integración de Alumnos con Necesidades Educativas Especiales*. Tercer informe. MEC. Madrid.
- VARIOS (1992): *Evaluación de la calidad de los Servicios del Programa de Termalismo Social*. Ministerio de Asuntos Sociales, INSERSO, Madrid.

REFERENCIAS BIBLIOGRÁFICAS

- ABARCA, P. (coord.) (1989): *La evaluación de Programas Educativos*. Escuela Española, Madrid.
- ALKIN, M. C. (1991): Evaluation theory development. En M. W. McLaughlin y D. C. Phillips (eds.): *Evaluation and education*. At Quarter Century, National Society Of Education, Chicago.
- ÁLVAREZ, M. y otros (1988): «Evaluación de programas de Orientación: La evaluación del contexto y del diseño». Ponencia presentada al *IV Seminario de Modelos de Investigación Educativa*. Santiago de Compostela.
- ALVIRA, F. (1991): *Metodología de la evaluación de programas*. CIS, Cuadernos Metodológicos nº 2, Madrid.
- ANGUERA, M. T. (1989): Innovaciones en la metodología de la evaluación. *Anales de Psicología*, nº 5, pp. 14-42.
- ANGUERA, M. T. (1990): Programas de intervención. ¿Hasta qué punto es factible la evaluación? *Revista de Investigación Educativa*, nº 16, pp. 77-93.
- ANGULO, J. F. (1988): «Evaluación de programas sociales: De la eficacia a la democracia». *Revista de Educación*, 286, pp. 193-207.
- ATO, M. y otros (1989): Evaluación de programas: Aspectos básicos. *Anales de psicología*, nº 5, pp. 3-12.
- BANK, A. y MORRIS, L. L. (1979): *A needs assessments kit: Guidelines and Resources for Educators*. Center the Study of Evaluation. University of California, Los Ángeles.
- BARTOLOMÉ, M. (1990): Evaluación y optimización de los diseños de intervención. *Revista de Investigación Educativa*, nº 16, pp. 39-60.
- BLANCHET, A. (1989): «Entrevistar» en A. Trognon y otros: *Técnicas de investigación en ciencias sociales*. Narcea, Madrid.
- BERK, R. y ROSSI, P. M. (1990): *Thinking about program evaluation*. Sage, Beverly Hills.
- BRUNET, L. (1987): La investigación evaluativa en el medio organizacional. En Y. Bordelau: *Modelos de investigación en recursos humanos* (pp. 122-149): Trillas, México.
- CAMPBELL, D. T. y STANLEY, J. C. (1973): *Diseños experimentales y cuasiexperimentales en la investigación social*. Amorrortu Editores, Buenos Aires (versión española de 1978).
- CARIDE, J. A. (1989): «De la evaluación de necesidades a la evaluación de programas sociales en el desarrollo comunitario». Ponencia inédita.
- COHEN, E. y FRANCO, R. (1989): *Evaluación de proyectos sociales*. G.E.L., Buenos Aires.
- COLÁS, P. (1992): Evaluación de programas en el estudio de cuestiones educativas. *Cuestiones Pedagógicas*, nº 4-5, pp. 79-91.
- COLÁS, P. y REBOLLO, M. A. (1993): *Evaluación de programas. Una guía práctica*. Editorial Kronos, Sevilla.
- COOK, T. D. y REICHARDT, CH. (1986): *Métodos cualitativos y cuantitativos en la investigación educativa*. Morata, Madrid.
- COOK, T. D. y SADISH, W. R. (1986): Program Evaluation: The worldly science. *Annual Review of Psychology*, 37, 193-231.
- CRONBACH, L. J. (1980): *Toward reform of program evaluation*. Jossey-Bass, San Francisco.
- DANIELS, M. H. y otros (1981): A meta-model for evaluating counseling programs. *Personnel and Guidance Journal*, 59, 578-582.
- DE LA ORDEN, A. (1985 a): Hacia una conceptualización del producto educativo. *Revista de Investigación Educativa*, nº 16, 61-76.
- DE LA ORDEN, A. (1985 b): Investigación evaluativa. En A. de la Orden: *Investigación Educativa*. Diccionario de Ciencias de la Educación (pp. 133-137): Anaya, Madrid.

- DE LA ORDEN, A. (1990): Evaluación de los efectos de los programas de intervención. *Revista de Investigación Educativa*, nº 16, pp. 61-76.
- DUNN, R. y DUNN, K. (1977): *Administrator's guide to new programs for faculty mangement and evaluation*. Parker Publ., New York.
- FERRÁNDEZ, A. (1990): Identificación de la evaluación en Educación de Adultos: las alternativas. *V Jornadas de Educación de Adultos*. Zaragoza.
- FINNEGAN, J. R. y otros (1988): Measuring and tracking education program implementation: The Minnesota Heart Health Program. *Health Educational Quarterly*, 16 (1), 77-90.
- FITZ-GIBBON, C. T. y MORRIS, L. L. (1978): *How to design a Program Evaluation*. Jossey-Bass, San Francisco.
- FLANAGAN, J. C. (1954): «The critical incident technique». *Psychological Bulletin*, 54, 327-358.
- FORNI, F. (1988): *Formulación y evaluación de proyectos de acción social*. Humanitas, Buenos Aires.
- GAIRÍN, J. (1991): *La evaluación de programas y cursos*. Documento policopiado. Bellaterra, Barcelona.
- GARANTO, J. (1989): Modelos de evaluación de programas educativos. En *La evaluación de Programas Educativos*. Escuela Española, Madrid, pp. 43-78.
- JUDD, C. M. y KENNY, D. A. (1981): *Estimating the effects of social interventios*. Cambridge Univ. Press, New York.
- LEITHWOOD, J. y MONTGOMERY, D. J. (1980): Evaluating program implementation. *Evaluation Review*, 4 (2), 193-214.
- LEVIN, H. M. (1985): *Cost-effectiveness*. Sage, Beverly Hills.
- MARTINEZ, A. y otros (1990): Evaluación de los programas de intervención en las instituciones de menores (Valencia): *Revista de Investigación Educativa*, nº 16, pp. 325-331.
- MASSONNAT, J. (1989): «Observar» en A. Trognon y otros: *Técnicas de investigación en ciencias sociales*. Narcea, Madrid.
- MAYER, R. R. (1985): *Policy and program planning*. Prentice Hall, Englewood Cliffs, N.J.
- POPHAM, W. J. (1980): *Problemas y técnicas de evaluación educativa*. Anaya, Madrid.
- POZO, M. T. y otros (1990): Evaluación de un programa de intervención en educación ambiental. La investigación-acción como estrategia de conservación del medio natural. *Revista de Investigación Educativa*, nº 16, pp. 309-315.
- RUTMAN, L. (1980): *Planning useful evaluations. Evaluability assessment*. Sage, Berverly Hills.
- SANZ ORO, R. (1990): *Evaluación de programas de orientación educativa*. Pirámide, Madrid.
- SELDIN, P. (1985): *Changing practices in faculty evaluation*. Jossey-Bass, San Francisco.
- SHIPMAN, S. (1989): General criteria for evaluating social programs. *Evaluation Practice*, 10 (1), 20-26.
- STUFFLEBEAM, D. L. y otros (1984): *Conducting Educational Needs Assessment*. Kluwer-Nighott Publ., Boston.
- STUFFLEBEAM, D. y SKINKFIELD, A. (1987): *Evaluación sistemática: Guía teórica y práctica*. Paidos-Mec, Madrid.
- STUFFLEBEAM, D. (1988): Normas para evaluadores. En R. Pascual (eds.): *La gestión educativa ante la innovación y el cambio*. Narcea, Madrid.
- STUFFLEBEAM, D. (1991): Professional standards a ethics for evaluators. En M. W. McLaughlin y D. C. Phillips (eds.): *Evaluation and education*. At Quarter Century, National Society Of Education, Chicago.
- TALMAGE, H. (1982): Evaluation of programs. En *Encyclopedia of Educational Research*. McHollan, New York.
- TEJADA, J. (1991): *La evaluación en la formación ocupacional*. Documento policopiado. Bellaterra, Barcelona.
- TEJEDOR, F. J. (1990): Perspectivas metodológicas del diagnóstico y evaluación de necesidades. *Revista de Investigación Educativa*, nº 16, pp. 15-38.

- VAN DE VEN, A. H. y FERRY, D. L. (1980): *Mesasuring and assessing organizations*. Wiley, New York.
- WEIS, C. H. (1982): *Investigación evaluativa. Métodos para determinar la eficiencia de los programas de acción social*. Trillas, México.
- WHOLEY, J. S. (1979): *Evaluation: Promise and Perfomance*. The Urban Institute, Washington.
- WITKIN, B. R. (1982): «New York Association for Learning Disabilities Project: Formative Evaluation-Fault tree Compnent. Final report» en K. G. Butler y otros *Evaluation of Sulffolk and Nassau County Chapter Projects*. Syracuse University.
- WITKIN, B. N. (1984): *Assessing needs in Educational and Social Programs*. Jossey-Bass Publ., London.
- WORTMAN, P. (1983): Evaluation research: A methodological perspective. *Annual Review of Psychology*, 34, 223-240.

PONENCIA III

MODELOS DE CONSTRUCCIÓN Y VALIDACIÓN DE INSTRUMENTOS DIAGNÓSTICOS

por

*Arturo de la Orden Hoz, José Luis Gaviria Soto,
Aurora Fuentes Vicente, Ángel Lázaro Martínez*
Universidad Complutense de Madrid

1. INTRODUCCIÓN

Esta ponencia se centra en las alternativas instrumentales en la evaluación diagnóstica de los alumnos en el contexto del sistema educativo. Independientemente de sus múltiples connotaciones, en este trabajo, identificamos el diagnóstico con la evaluación tendente a la identificación de los procesos mentales que subyacen el funcionamiento cognitivo de los alumnos con la finalidad de determinar el patrón de las representaciones cognitivas de un sujeto y el estadio de su conocimiento y destreza en relación con la competencia total en un campo de instrucción determinado.

Las pruebas existentes hasta ahora han permitido situar a los sujetos en un conjunto de capacidad o competencia global, pero no analizar la estructura cognitiva propia de un sujeto en un determinado estadio (diagnóstico): Se precisan nuevos test o, más precisamente, nuevos tipos de ítems. La adecuación de un ítem para ser aplicado en un momento determinado dependerá del conocimiento que se tenga del sujeto. Este conocimiento permite elegir en cada momento el ítem potencialmente más informativo. Obviamente, para determinar cuál es la naturaleza de los ítems que se precisa aplicar para incrementar el conocimiento acerca del patrón procesual del sujeto, es necesario partir de un modelo cognitivo de alumno o aprendiz, de modo análogo a como se procede en el diagnóstico médico, partiendo de un conocimiento de la relación causal entre enfermedad y síntomas como guía para decidir el tipo de pruebas necesarias para detectar ciertos síntomas.

En resumen, el desarrollo de una nueva generación de tests diagnósticos no tiende simplemente a conseguir los mismos fines que las pruebas tradicionales con

mayor perfección o eficiencia, sino que, a partir de una reconceptualización del propósito de los tests en general y de la propia función evaluativa, apunta a obtener una información diferente de cada sujeto para responder a las exigencias de mejora de la instrucción. Concretando, el diagnóstico a través de los nuevos tests es concebido como un medio para desarrollar un retrato muy elaborado de un individuo en orden a facilitar la toma de decisiones en un proceso de aprendizaje en marcha (Frederiksen, Mislevy y Bejar, 1993).

Para desarrollar una evaluación diagnóstica de este tipo, probablemente (Shepard, 1991) sería necesario girar 180 grados lo que Resnik y Resnik (1989) llamaron presunciones de descomponibilidad y descontextualización de las teorías tradicionales del aprendizaje. Los tests no deben pedir la demostración de pequeñas destrezas discretas practicadas aisladamente, sino que deben convertirse en instrumentos más ambiciosos que intenten detectar las representaciones mentales que los estudiantes tienen de ideas importantes y su facilidad de aplicarlas a la resolución de problemas nuevos.

¿En qué grado, pues, los tests actualmente en uso generalizado y las teorías que los subyacen pueden responder a estas exigencias diagnósticas?

En un informe de investigación, Mislevy (1993a), uno de los más productivos académicos del Educational Testing Service, afirma sin rubor que resulta solamente un poco exagerado describir la teoría de los tests, hoy dominante en la medida y el diagnóstico educativos, como la aplicación de la «estadística del siglo XX a la psicología del siglo XIX». Así, procedimientos sofisticados de estimación y avances teóricos sobre modelamiento de variables latentes, por ejemplo, se aplican a modelos psicológicos que intentan explicar la capacidad para resolver problemas en términos de una variable continua única. Aunque tomadas literalmente estas palabras van más allá de lo que parece razonable inferir a partir de un análisis de la realidad, resulta evidente una cierta incongruencia entre las pruebas convencionales y las nuevas finalidades diagnósticas.

El problema básico es que la visión de las capacidades humanas implícita en la teoría convencional de los tests —tanto en la teoría clásica (respuesta verdadera) como en la IRT— no parece compatible con la visión que emerge con gran fuerza y rapidez de la moderna psicología cognitiva y educacional. Los alumnos incrementan su competencia no por la simple acumulación de nuevos datos y destrezas, sino por la reconfiguración de sus estructuras de conocimiento, por la automatización de procedimientos y reorganización de la información para reducir peso en la memoria y por el desarrollo de estrategias y modelos que les indican cuándo y cómo los datos y las destrezas son relevantes. Los tipos de observaciones y los patrones en los datos, que reflejan la forma en que los alumnos piensan, realizan y aprenden, no pueden acomodarse a los métodos y modelos tradicionales.

La base, pues, de una nueva teoría de los tests sería la aplicación de los modernos métodos estadísticos a modelos psicológicos también modernos.

En realidad, la psicometría es la historia de un proceso de formalización. Podemos decir que hay una continuidad básica en el proceso de modernizar la conducta

de un sujeto al responder a un test. Las tendencias psicométricas actuales intentan superar las limitaciones detectadas en las fases previas del proceso de modelizar alumnos. Estas fases pueden concretarse en la teoría clásica de los tests y en la teoría de respuesta al ítem, que constituyen los antecedentes de las corrientes más recientes.

En la teoría clásica, el test es considerado como un estímulo único, cuya respuesta observada está incluida en el modelo solamente como una única cantidad, X . Ésta es la única magnitud observable, y el modelo no prejuzga cómo se obtiene esta medida. De hecho, este modelo representa una respuesta observable y cuantificable de los sujetos a cualquier tipo de estímulo, aunque lo más común sea la consideración de X como proveniente de la respuesta a un test de lápiz y papel.

La fiabilidad es parte del núcleo fundamental de esta teoría. Se define como la proporción de varianza observada que es varianza verdadera. Como tal, la fiabilidad, probablemente la más importante característica técnica en esta teoría, es una propiedad del test. Para su definición operativa es preciso la definición de las condiciones de paralelismo. Sólo en el caso de que dichas condiciones se cumplan es posible poner en relación la fiabilidad, una magnitud inobservable, con alguna cantidad observable.

Una nota peculiar de esta teoría, consecuencia de la definición del modelo, es que algo tan común en el trabajo ordinario de los constructores de tests, como los ítems, no está incluido en el mismo. Por esa misma razón el conjunto de técnicas y parámetros con los que ordinariamente se opera cuando se trabaja con los ítems, resulta ser un añadido que no tiene relación formal con el núcleo de la teoría. Así, resulta que las características técnicas de los ítems son cantidades referidas siempre a una muestra específica de sujetos y a un conjunto concreto de ítems. Por esto, el cambio tanto de grupo de referencia, como de ítems determinados con los que uno dado se presenta, hace que cambien sus características técnicas como dificultad y discriminación. En particular la discriminación de un ítem depende de cómo son los restantes ítems del test. La dificultad es función de los sujetos que lo han contestado. La dificultad, se define como la simple proporción de respuestas correctas dada al ítem, y está por eso mismo definido en una escala distinta de aquella en la que se presentan las puntuaciones de los sujetos. No hay, en definitiva, ninguna explicación del origen psicológico de la dificultad, y ni siquiera está relacionada con la magnitud que el test mide.

Esto nos lleva a hablar del modelo psicológico latente en la Teoría clásica de los tests. No puede hablarse de un único modelo psicológico. Existen razones históricas que han llevado a una determinada formulación, y varios influjos de corrientes psicológicas dominantes en cada período, como el funcionalismo y las teorías factoriales, y el conductismo.

La teoría de respuesta al ítem implica un avance en varios sentidos con relación a la teoría clásica. En primer lugar, el estímulo cuya respuesta se quiere modernizar ya no es un test, sino más bien un ítem. De hecho esta teoría tampoco prejuzga como se obtiene la puntuación de un ítem, aunque lo más frecuente sea referirse a

ítems de opción múltiple en tests de lápiz y papel. En estos modelos, lo esencial es la curva característica del ítem, que pone en relación la puntuación de un sujeto en una variable latente, con la probabilidad de responder correctamente. Los distintos modelos de respuesta al ítem tienen en general la forma

$$P(u_{ij} = 1 | \theta_j) = \varphi(a_i(\theta_j - b_i))$$

donde u_{ij} es la puntuación del sujeto j en el ítem i , siendo 1 si la respuesta es correcta, y 0 en caso contrario, a_j y b_j son, respectivamente, los parámetros de discriminación y dificultad del ítem j .

Una primera ventaja resulta de la integración entre las técnicas de análisis de ítems y las características técnicas de los tests. De hecho en esta teoría los tests son agregados de ítems cuyas características técnicas son acumulativas. La más importante característica es la invarianza de los parámetros de los ítems respecto a las muestras y al resto de los ítems.

Con relación a la dificultad, en la Teoría de respuesta al ítem este importante parámetro sitúa al ítem en la misma escala que las puntuaciones de los sujetos. Aquí se define ya la dificultad en relación a la magnitud latente que queremos medir. La dificultad de un ítem viene dada por aquél punto de la escala en el que se produce la inflexión de la curva característica. Esta definición resulta más clara si la referimos a un modelo de uno o dos parámetros. En ese caso la dificultad de un ítem viene dada por la puntuación para la que la probabilidad de responder correctamente es 0.5. Aunque no hay una explicación causal de la dificultad, ya tenemos al menos una definición más psicológica de la misma. Mientras en la teoría clásica la dificultad es simplemente una proporción, una cantidad que es totalmente contingente respecto a la muestra, aquí su definición tiene una mejor interpretación, por una parte, mientras que por otra nos permite situar a los ítems en el mismo continuo que los sujetos.

Los modelos iniciales de IRT no suponen ninguna ruptura con respecto a las ideas psicológicas dominantes en la psicometría anterior, pero su propia formulación supone la posibilidad de superar ese marco. Y lo que lo hace posible es que, mientras que en la teoría clásica los supuestos iniciales y enunciados no siempre han estado claramente explicitados, en la IRT los axiomas han sido formulados de forma clara y meridiana desde un primer momento (Lord & Novick, 1968): El supuesto de unidimensionalidad no es más que un caso particular del supuesto más general denominado de «**independencia local**» o «**independencia condicional**». Así expresado este principio supone que las correlaciones entre los ítems quedan completamente explicadas por el rasgo o variable latente a todos los ítems. Dicho de otra forma, si obtenemos la correlación entre todos los posibles pares de ítems, parcializando el rasgo latente, esas correlaciones son cero. Naturalmente lo que se propugna no es que los ítems de un test no están correlacionados entre sí, sino que lo único que explica que estén correlacionados es la dimensión única que todos miden en común. En definitiva el supuesto de independencia local equivale a conceder carta

de naturaleza causal a las variables condicionantes. Dados dos fenómenos concomitantes, afirmar su independencia condicionada a los valores de cierta variable, equivale a decir que esa variable es la única causa que explica aquellos efectos. La importancia de este supuesto en la causalidad en general ha sido destacada por Pearl, 1988:

«...la independencia condicional no es una gracia de la naturaleza por la que debemos esperar pasivamente, sino más bien una necesidad psicológica que satisfacemos activamente organizando nuestro conocimiento en una forma específica. Una importante herramienta en tal organización es la identificación de variables intermedias que inducen independencia condicional entre [variables] observables; si tales variables no están en nuestro vocabulario, las creamos. En diagnóstico médico, por ejemplo, cuando algunos síntomas se influyen mutuamente de forma directa, la profesión médica busca un nombre para esa interacción (es decir, “síndrome”, “complicación”, “estado patológico”) y la trata como una nueva variable auxiliar que induce independencia condicional; la dependencia entre dos sistemas interactuantes se atribuye completamente a las dependencias de cada uno con la variable auxiliar.» (p. 44)

La inclusión por tanto de este supuesto confiere a la teoría de respuesta al ítem una cualidad de explicación causal que no estaba presente en los modelos anteriores.

En definitiva, este supuesto, que es el axioma central de la teoría, nos permite romper con el marco inicial de referencia, ya que podemos pensar en varios rasgos latentes, que en conjunto cumplirían con el supuesto de independencia local. De esta forma se abre la puerta a modelos multidimensionales (McDonald, 1985 para un planteamiento general; Wang, 1985; Wang, 1987 y Reckase, 1979, para el estudio de las consecuencias de ajustar modelos unidimensionales a datos multidimensionales): Y lo más importante es que ahora tenemos unas condiciones que permiten tratar de generar nuevos modelos o adaptar los existentes a las condiciones que impongan las aplicaciones prácticas. Y esto ha ocurrido abundantemente, como veremos. Inicialmente estos modelos, que podríamos llamar de segunda generación, se formulan, la mayoría de las ocasiones, sin pretender modificar los supuestos psicológicos de base. Pero si *es posible generar modelos que se adaptan a exigencias prácticas* de distinto tipo, *se está abriendo la puerta a la posibilidad de hacer lo mismo con relación a las ideas psicológicas respecto al aprendizaje y la cognición.*

Uno de los primeros modelos que podríamos llamar «no-estándar» fue el desarrollado por Bock (1972): La idea original era que en un test con ítems de opción múltiple, la información de que disponíamos era no sólo la corrección o incorrección de la respuesta del sujeto, sino también cuál era exactamente la opción elegida. Así considerado el ítem ya no es dicotómico (1=correcto, 0=incorrecto), sino politómico (1,2,3,4...etc):

En la misma línea de especificar modelos para otros tipos de respuestas, Samejima (1969) produjo un modelo, o más bien las condiciones para generarlo, para ítems con respuesta graduada. En estos ítems las respuestas están ordenadas de mayor a menor corrección. El modelo especifica la curva (Trace line) correspondiente a la probabilidad de responder la opción k u otra superior. Ésta es la función. Esta función puede ser cualquiera de los modelos estándar para ítems dicotómicos, logísticos, de ojiva normal o cualquier otra variante. La curva característica de cada opción, aquella que relaciona la probabilidad de responder exactamente la opción k , viene dada por la diferencia entre dos trace lines

$$T_k(\theta) = T_k^*(\theta) - T_{k+1}^*(\theta).$$

Estos modelos permiten aplicaciones distintas de los tests de conocimientos ordinarios. Se ha comprobado, (Koch, 1983), que el modelo de repuesta graduada de Samejima es muy apropiado para analizar escalas actitudinales tipo Likert. La principal ventaja es que aprovecha al máximo la información aportada por cada ítem, con lo que el error de medida es mucho más pequeño.

Masters (1982), desarrolló otro modelo, denominado «Partial Credit Model». Se trata de otro modelo para respuestas graduadas, Mientras que el de Samejima pertenece al tipo de modelos de diferencias, según la taxonomía de Thissen y Steinberg (1986), el de Masters pertenece a la clase de división por el total según esa misma clasificación¹.

Masters y Wright (1984) formularon un modelo más general, que englobaba como casos particulares a otros muchos. El modelo dicotómico, el «Partial Credit», y otros como el «Rating Scale», o el de ensayos binomiales que no hemos mencionado aquí, son casos particulares de este modelo general.

Recientemente, Muraki (1992), ha desarrollado otra generalización de este modelo, denominada, «Generalized Partial Credit Model», junto con el procedimiento de estimación de parámetros en un programa denominado «Parscale».

Se han formulado muchos otros modelos psicométricos de características similares a los mencionados, aunque no los abordaremos aquí.

La nueva psicometría, cuyos antecedentes inmediatos los constituyen algunos modelos como los que acabamos de mencionar, parte de la constatación de que los fundamentos psicológicos utilizados en los modelos psicométricos habituales no son suficientes para representar la complejidad tanto de la estructura de conocimientos de los sujetos como de los cambios en esa estructura que el proceso de aprendizaje entraña (Véase por ejemplo Sternberg, 1984): Elementos éstos en los que la contemporánea psicología cognitiva ha centrado su atención.

La psicología cognitiva pone su interés en el estudio de la **adquisición, organización y representación del conocimiento**. Ha tomado como modelo central el de

1 A ese mismo grupo pertenece el modelo de Bock para respuestas nominales.

procesamiento de la información. Debido a esto, en palabras de SNOW y LOHMAN (1989),

El computador, entonces, se ha convertido en una metáfora para la teoría y un medio para su realización y evaluación.

Los temas clásicos de la psicología cognitiva son el de el reconocimiento de formas, la percepción y la atención, el aprendizaje, la memoria, el razonamiento, la resolución de problemas, el pensamiento, la comprensión y producción de lenguaje y la representación de conocimientos, tanto de naturaleza declarativa como procedimental. De todos estos temas probablemente el más recurrente en la literatura es la **resolución de problemas**, ya que es, no sólo una función mental, sino una actividad que implica a todas las demás funciones mentales. Para Elshout (1985), la resolución de problemas no es un dominio especial, sino una preocupación fundamental de todos los dominios. Pues bien, la resolución de problemas liga íntimamente a la psicología cognitiva con la psicometría. O dicho de otra forma, hace del proceso de respuesta a los ítems de un tests objeto de interés de la psicología cognitiva, pues eso es en sí mismo un proceso de resolución de problemas.

Desde un punto de vista cognitivista, los tests tradicionales de rendimiento, fallan en la evaluación de la arquitectura de las funciones mentales básicas, dado que su criterio de construcción es el contenido del dominio, bajo los supuestos de una cierta homogeneidad del mismo, y de una continuidad en el proceso de adquisición de los conocimientos implicados en ese dominio. Los modelos psicométricos elaborados para dichos tests tratan de situar a los sujetos en un continuo. La puntuación que se asigna al sujeto no tiene un claro significado psicológico. Se trata en realidad de un artificio, de un resumen conveniente que refleja una combinación compleja de destrezas, habilidades, conocimientos y otros componentes cognitivos.

Lo que realmente falta en la conceptualización en que se apoya la teoría clásica de los tests son modelos para representar cómo conocen los individuos aquéllo que conocen, cómo hacen aquéllo que hacen y la forma en que incrementan la capacidad de conocer y de hacer. La cuestión no es que los antiguos modelos sobre el nivel de capacidad sean falsos y los nuevos modelos verdaderos, sino que diferentes modelos son útiles para diferentes propósitos y hemos de desarrollar una teoría general para razonar a partir de las observaciones hacia más amplios modelos de habilidad congruentes con la investigación sobre cognición y capaces de resolver cuestiones educativas a partir de la aplicación de esta teoría.

En resumen, siguiendo a Mislevy (1990), puede afirmarse que, desde el punto de vista diagnóstico, las limitaciones de los tests tradicionales quedan bien reflejados en la frase de Glaser, Lesgold y Lajoie (1987) cuando afirman que los tests (usuales) pueden (en cierta medida) predecir el fallo de ciertos alumnos sin una comprensión de lo que causa el éxito; pero, la intervención para prevenir el fallo e incrementar la competencia requiere una más profunda comprensión.

Según Snow y Lohman (1993), ahora ya es posible identificar algunas de las

destrezas procesuales componentes de la capacidad medida por los tests: codificación de estímulos, comparación de facetas, inducción de reglas, aplicación de reglas, justificación de respuestas, son ejemplos de tales procesos.

Es decir, el análisis cognitivo de los tests proporciona una explicación psicológica de los procesos cognitivos que se producen al enfrentarse con la resolución de las tareas planteadas en los ítems, y por tanto, de la dificultad de los propios ítems. Ésta proviene de varios atributos que el ítem adquiere en el proceso de su construcción. Esa dificultad depende de los componentes cognitivos que el ítem implica. Desde un punto de vista cognitivo nos interesa saber por qué se ha producido determinada respuesta, tanto sea correcta como incorrecta. Se trata de desvelar la caja negra que antes constituía un ítem o un test. Ahora no nos interesa sólo saber si se respondió correcta o incorrectamente a un determinado ítem, que desde un punto de vista cognitivo es una tarea a resolver, sino cuáles son las razones psicológicas, cuál es la arquitectura de conocimientos y destrezas específicas del sujeto, y cuáles los requerimientos cognitivos del ítem que explican ese resultado.

Butterfield y otros (1985), mostraron cómo los análisis cognitivos pueden conducir a procedimientos para la generación computerizada de ítems diseñados para controlar las diversas fuentes de dificultad. El resultado será diseños de pruebas que ayuden a diagnosticar las incapacidades asociadas con tipos de dificultad determinados. Sin embargo, el modelo psicométrico que se precisa para tales tests debe acomodarse a un proceso diagnóstico adaptativo en el que la sucesión de ítems es configurada para cada persona en cada campo específico.

El análisis cognitivo, (Snow y Lohman, 1993), también ha pretendido comprender el conocimiento proposicional (declarativo) y procesual adquirido por los alumnos durante la instrucción formal. Un producto típico de la psicología cognitiva es una visión más rica de la organización del conocimiento tanto en su adquisición como en su estructura. Esta visión distingue diversas fases de conocimiento y de adquisición de destrezas. La investigación apunta a diferentes formas de estructura de conocimiento producida por el aprendizaje. Entre estas formas cabe citar redes semánticas, «schemata», prototipos, imágenes y modelos mentales. Estas estructuras de conocimiento se forman no únicamente durante el aprendizaje, sino que se utilizan también para razonar sobre lo que está siendo aprendido, para recordar lo que ha sido aprendido y para resolver problemas.

Dado que diferentes clases de estructuras de conocimiento producen diferentes patrones de respuestas y tipos particulares de errores, los tests deben diseñarse para detectar estas diferencias.

Todo ello, además de evaluar la trayectoria del discente desde estados anteriores de conocimiento y destreza hasta la total competencia en un campo a través de la instrucción.

En esta línea, Haertel y Wiley (1993), frente a las tecnologías de las pruebas de elección múltiple, e incluso de la puntuación dicotómica de respuestas abiertas, apuntan a la necesidad de incrementar la atención a las transiciones en el aprendizaje entre estados de realización de tareas. Un análisis más minucioso de la realización

de tareas complejas centra la atención en secuencias de estados de realización que conducen, finalmente, a estados terminales que han sido las dianas principales de análisis anteriores. El nuevo marco de referencia yuxtapone las experiencias de aprendizaje en las que toma parte el individuo y los estados de capacidad con que entra en tales experiencias. La interacción de modelos de capacidad con experiencias de aprendizaje resulta en un nuevo aprendizaje, es decir, se procede a una transición a nuevos patrones de capacidad.

Intentando recoger estas ideas y dar parcialmente respuesta a alguna de ellas se han realizado algunos progresos en la teoría de la medida. Algunos ejemplos muy interesantes son el modelo de Fischer (1973) (véase también Fischer y Formann, 1982) (Linear latent trait model), el de Embretson (1984) (General latent trait model), generalización del formulado en Whitely (ahora Embretson) (1980) (Multi-component latent trait model): Estos son ejemplos de cómo una teoría de los tests diferente puede dar respuesta a una teoría de la complejidad de las destrezas procesuales componentes de la inteligencia.

En el nuevo enfoque se pretende dar respuesta global a todas estas exigencias, recogiendo algunas de las formulaciones mencionadas, e integrándolas en una nueva teoría general de los tests, cuya característica más sobresaliente es la conciencia de la voluntad expresa de incorporar los conocimientos psicológicos provenientes del análisis cognitivo con el objeto de proporcionar una herramienta útil para mejor integración del diagnóstico y el proceso de aprendizaje.

En esta línea, la naturaleza y el papel de los ítems son concebidos de modo diferente. En la concepción evaluativa tradicional, los ítems son estímulos cuyas características psicométricas conocidas y expresadas a través de sus parámetros nos permiten, por medio de las respuestas que elicitamos de los sujetos, situar a estos en el continuo hipotetizado. No existe una forma sistemática de generación de tales ítems, excepto en lo que se refiere al conjunto de contenidos y destrezas que constituyen el referente del continuo Y salvo la de que estén debidamente representados en el test los distintos subconjuntos implicados, no existe una regla para determinar cuál debe ser el contenido específico de un ítem dado.

En el nuevo enfoque, **el propio test se convierte en un diseño de observación de tareas**, en cuya realización se manifiestan las habilidades puestas en efecto por los sujetos. Los ítems por tanto tienen la forma más adecuada para que se objetiven aquellas acciones que queremos observar. El contenido del ítem o tarea a realizar viene determinado por las necesidades diagnósticas en un momento dado del proceso de evaluación.

En cuanto a las características psicométricas de los ítems, en el marco tradicional la más importante de todas es la dificultad. En la IRT ésta viene dada por el lugar que ocupa el propio ítem en el continuo, mientras que en la TCT está definida en una escala distinta de aquella en la que están situados los sujetos. No existe una explicación psicológica de la dificultad. En este sentido la dificultad de un ítem es una información conseguida siempre a posteriori. Es una magnitud que aparece como un atributo primario del ítem, que sólo puede ser empírica, o inductivamente

conocido. De hecho, los procesos de estimación de parámetros no dejan de ser procesos inductivos, generalizaciones empíricas que acarrearán consigo el error y la incertidumbre.

En esta nueva etapa de la teoría de los tests a la que nos estamos refiriendo, las características psicométricas de los ítems reciben una explicación causal. La dificultad de un ítem es una información a priori, que depende causalmente de las instancias cognitivas que intervienen en el mismo. El tipo de componentes cognitivos implicados en la resolución del ítem determinan la dificultad del mismo. Por lo tanto, es una magnitud que se deduce, y dado que siempre es una hipótesis, el proceso de estimación de parámetros es más bien un experimento en el que se lleva a cabo una contrastación de teorías. Por esto mismo, la diferencia entre la dificultad propugnada y la dificultad obtenida es una medida de la discrepancia entre nuestro modelo cognitivo y la realidad. En este modelo, el diagnóstico es un proceso constante de construcción de teoría.

Esto nos lleva a hacer algunas consideraciones importantes acerca de la validez. En realidad no se plantea la validez de constructo de forma radicalmente distinta. En esencia la validez de constructo implica la existencia de una explicación teórica de las respuestas obtenidas a un test. Pero mientras que en las teorías estándar de los tests se comprueba a través de la consistencia de las respuestas, en la consideración o enfoque cognitivo no basta con eso. Es preciso también que se dé una explicación de la dificultad de los ítems. Y eso exige un modelo teórico más elaborado, lo que en definitiva habla no sólo de una profundización en la exigencia de la validez de constructo, sino que además ofrece una nueva vía para la comprobación de la misma a través de los contrastes referidos a la dificultad.

No sólo en el nivel de los ítems hay un cambio de perspectiva. Para la medida de procesos cognitivos, la nueva teoría distingue dos aspectos de los tests. Por un lado, los tests pueden ser considerados como diseños de observación. Y por otro, las puntuaciones de los tests pueden ser consideradas como productos de diferentes modelos de medida.

El diseño de observación describe los ítems del test, su organización y el tipo de respuesta requerido. El propósito del diseño de observación es estructurar las observaciones de modo que a partir de ellas puedan hacerse inferencias defendibles acerca de constructos teóricos. La construcción de un test requiere compromiso con un diseño particular de observación. Por ejemplo, asume que cierto modelo de proceso subyace la realización de una tarea específica y que este modelo de proceso especifica una serie de componentes procesuales independientes. Dados estos supuestos las observaciones deben organizarse de tal forma que sea posible una prueba de la independencia de estos componentes procesuales.

El segundo aspecto de la prueba, es decir, el diseño de medida que se refiere al procedimiento para asignar un valor único a un objeto de medida. El análisis de proceso de las tareas cognitivas hace posible especificar múltiples objetos de medida, por ejemplo, la realización global de la tarea así como su nivel de competencia en cualquiera de los componentes procesuales. Para cada uno de estos objetos de

medida se utiliza un modelo que especifica las reglas que han de ser utilizadas para puntuar, clasificar o combinar objetos de observación. Por ejemplo, en la psicometría tradicional podía obtenerse una puntuación única del sujeto promediando la realización en todos los ítems. Esto es coherente con un modelo de medida que relega la varianza en la realización en los diversos ítems al término de error. En el enfoque de procesamiento de la información dar cuenta de la tarea ejecutada normalmente conduce a un conjunto más complejo de modelos de medida en que se estiman las puntuaciones para los componentes procesuales que los sujetos utilizan para generar las respuestas o elegir entre opciones de respuesta.

2. LA EVALUACIÓN DIAGNÓSTICA EN EL MARCO DE LA PSICOLOGÍA COGNITIVA

Si debe existir una característica distintiva del diagnóstico en contraposición con la evaluación tradicional, es precisamente la necesidad de ligar estrechamente el proceso de la evaluación diagnóstica y la instrucción. Un test diagnóstico debe ser útil para la instrucción. Debe proporcionar información instructivamente útil.

Los tests han dejado de ser casi exclusivamente un modo de medir la competencia global para la selección y evaluación de los sujetos y ha pasado a ser un modo de recoger información para ayudar en el proceso de aprendizaje y en la instrucción, en definitiva para ayudar a la toma de decisiones educativas.

Los nuevos tests deben responder a los retos que se les plantean. Es necesario disponer de instrumentos y estrategias que permitan conocer lo mejor posible el proceso de comprensión de los sujetos: qué han logrado en ese proceso, qué deficiencias sufren, qué aspectos les facilitan un aprendizaje posterior, qué aspectos lo retardan, qué malas interpretaciones y errores obstaculizan su avance, etc. Todo ello con el fin de lograr un verdadero aprendizaje y un adecuado nivel de competencia en un dominio determinado.

Esto exige una nueva teoría de los tests que incorpore los conocimientos proporcionados por la psicología cognitiva. Es preciso lograr la integración de la teoría psicométrica y la teoría cognitiva. El resultado de dicha integración tiene que ser capaz de dar explicación de las respuestas observables de los sujetos a las tareas que se les planteen, indicador falible, pero único disponible. La nueva generación de modelos psicométricos

... permiten incorporar la teoría cognitiva directamente en el diseño de los tests (Embreston, 1993, p. 148)

Los modelos basados en el procesamiento de la información que se proponen describir cómo los componentes procesuales se organizan en estrategias, deben de alguna manera estimar la acción de cada proceso postulado. Sin embargo, los procesos mentales son inobservables. Deben ser inferidos a partir de patrones de respuestas a lo largo de un conjunto de ítems que se supone requieren diferente

cantidad de diferentes procesos. Para cumplir esta misión han sido propuestos diferentes diseños.

La evaluación diagnóstica considerada en el conjunto de todas sus fases y procedimientos, es esencialmente un proceso de inferencia. El proceso de diagnóstico consiste en establecer cuál es la arquitectura cognitiva del sujeto que aprende a partir de las observaciones realizadas en el diseño de observación, ya sea un test, una recopilación de realizaciones personales, un contexto de ejecución efectiva de una tarea, o una simulación. Pero en cualquier caso diagnosticar supone construir hipótesis sobre estados cognitivos latentes del sujeto, y por ende no observables, a partir de realizaciones observables, incluyendo en éstas tanto conductas como, de una forma ligeramente más indirecta, sus resultados. De esta forma el diagnóstico es inferencia, en el sentido dado por Mislevy,

“Inferencia es razonar desde lo que conocemos y lo que observamos hasta explicaciones, conclusiones o predicciones». (Mislevy, 1993b, p. 1)

La teoría de la medición educativa se convierte en este contexto en una herramienta al servicio del diagnóstico, siendo éste un proceso continuado de elaboración de hipótesis, extracción de evidencia empírica para contrastarlas, y nueva formulación de las mismas. La teoría de la medida desempeña en el proceso de diagnóstico el mismo papel que la estadística desempeña en el contexto de la investigación educativa, lo que refuerza la idea de la naturaleza experimental del proceso de diagnóstico. En este sentido puede entenderse la afirmación de Mislevy,

La teoría de los tests es una maquinaria estadística para extraer hipótesis acerca de la competencia de los estudiantes a partir de su conducta, según una determinada concepción particular de lo que la competencia significa en ese dominio. (Mislevy, 1993c)

La construcción de un instrumento diagnóstico, desde esta nueva perspectiva, podemos afirmar que supone el enfrentamiento con tres problemas básicos (Mislevy, 1991):

1. La formulación del modelo.
2. La construcción de las tareas cognitivas a observar.
3. La inferencia a partir de los datos.

2.1. La formulación del modelo

Los datos no son lo mismo que la evidencia. Ésta sólo existe en el marco de una hipótesis. El modelo es la hipótesis que hace que los datos se conviertan en evidencia empírica. Es el marco conceptual en el que las observaciones recogidas tienen sentido dentro de determinada teoría del aprendizaje.

Formular el modelo consiste en determinar los estados o niveles de comprensión

en un dominio de aprendizaje. Estos estados quedan recogidos en un vector-variable, por ejemplo η . Ese vector representa los parámetros del estudiante. Es una representación simplificada. Los parámetros del modelo especificado deben reflejar diferencias clave en la comprensión de ese dominio, cuáles son los conceptos importantes y cuáles las vías oportunas para entenderlos. Deben recoger diferencias fundamentales en la comprensión del mismo. Tales parámetros deben destacar el desarrollo de la comprensión del sujeto, por lo que en su determinación se ha de tener en cuenta tanto el contenido y estructura de tal dominio como la psicología del aprendizaje correspondiente. Sea cual sea su naturaleza métrica, (no ordenados, parcialmente ordenados, completamente ordenados), deben cumplir la condición señalada: captar distinciones esenciales en la descripción del nivel de comprensión de los sujetos.

Los mapas o modelos cognitivos son específicos de cada dominio, lo que plantea el problema de la gran variedad de los mismos en cuanto a su nivel de estructuración. Pues bien, dado un determinado dominio, ¿A qué nivel de «granularidad» con qué nivel de detalle debe representarse la estructura cognitiva del alumno? Una solución consiste en representar los modelos cognitivos a aquél nivel al que puede implementarse la **instrucción**.

La expresión de Greeno (1976) aclara bien este criterio:

«No es fundamental distinguir entre modelos que difieran en procesos carentes de implicaciones importantes para la calidad del trabajo académico del estudiante o para progresar en su conocimiento o comprensión»

Lógicamente un modelo de alumno simplifica la realidad. Pero mientras que en la investigación psicológica esto tiene connotaciones relativas a la validez de la teoría, **en el diagnóstico la plausibilidad del modelo tiene que ver sobre todo con su utilidad**. ¿Recoge el modelo las características necesarias para tomar una decisión instructiva adecuada? Esta pregunta es la que determina, como ya se ha señalado, el umbral de definición del modelo. Ir más allá supondría recoger detalles anecdóticos o estériles. Quedarse más acá supone perfilar modelos demasiado toscos para ser útiles.

Un modelo de alumno es una descripción simplificada de ciertos aspectos seleccionados de una variedad infinita de destrezas y conocimientos que caracterizan a los verdaderos estudiantes. Dependiendo de la finalidad, uno puede distinguir desde uno a cientos de aspectos. Podrían expresarse en términos de números, categorías, o alguna combinación.»
(Mislevy, 1993c)

La formulación del componente cognitivo de los nuevos modelos de medida implica una metodología de análisis del dominio de conocimientos que se quiere modelizar. Esta metodología aborda dos problemas importantes: la **elicitación de**

conocimientos de los sujetos implicados en las tareas relevantes, y la **representación** de la organización y estructura del mismo.

Hay tres importantes tendencias en la elicitación del conocimiento de un dominio:

- Análisis de protocolos verbales durante la acción. (Newell y Simon, 1972): Consiste en grabar el relato que el sujeto hace de la acción al mismo tiempo que la ejecuta.
- Entrenamiento (Coaching): (Gelman y Gallistel, 1978): El sujeto va explicándole al observador cómo tiene que hacer para que sea éste último quien desarrolle la tarea.
- Análisis de la comunicación ordinaria de trabajo dentro de un equipo. (Orasanu y Fischer, 1992):

En cuanto a la representación de los conocimientos hay dos líneas de investigación:

- Representaciones computacionales de arquitecturas cognitivas. Se trata de codificar un programa de ordenador que lleve a cabo las tareas representadas en un entorno simbólico. Tienen la ventaja de ser representaciones explícitas y completas. Su inconveniente es el costo elevado y la dificultad de realizar representaciones de dominios poco estructurados o altamente declarativos.
- Identificación de diferencias en componentes cualitativos del conocimiento. Consiste en identificar las diferencias cualitativas en el conocimiento que van asociadas con el éxito en la realización de tareas. Para ello se realizan comparaciones entre esquemas desarrollados por novatos y por expertos.

Los modelos a que nos referimos aquí deben incluir tanto un conjunto de parámetros que reflejen el esquema mental que queremos representar, producto del análisis cognitivo del dominio de conocimientos (η), como una estructura que simbolice la relación de probabilidad condicional de ese esquema con las tareas en que se pone de manifiesto, $P(x|\eta)$: Las representaciones que deben construirse en este contexto deben contar con dos elementos: una figuración razonable de la complejidad del proceso cognitivo y un mecanismo adecuado para tratar la incertidumbre derivada del hecho de que no hay una relación unívoca entre los desempeños de los alumnos en las tareas observadas y sus estados cognitivos latentes.

Y para dicha caracterización y diferenciación de parámetros básicos, es fácil de entender la ayuda que proporcionan los avances en el análisis de las estructuras de conocimiento que hemos mencionado, así como en el análisis de la estructura fina de las habilidades que condicionan la ejecución de los sujetos y de los procesos por los que se adquieren.

Son buen ejemplo en este sentido los trabajos que sugieren «esquemas» de esa conceptualización, los que proponen pautas para evaluarlos (Marshall 1990, 1993), o la valiosa aportación de Haertel y Wiley (1993): Estos últimos autores desglosan las características de las estructuras de habilidades latentes y tareas observadas en que se proyectan y, lo que es más importante, tratan explícitamente las diferentes relaciones que se pueden dar entre unas y otras.

Como consecuencia de lo anterior podemos afirmar que el proceso de construcción de un modelo de alumno tiene dos fases:

- **Definición.** Consiste en determinar qué elementos forman parte del modelo y cuales son sus relaciones estructurales. Entre los elementos deben incluirse necesariamente un conjunto de tareas observables y un conjunto de estados no observables. El modelo debe especificar cómo los segundos determinan los primeros. Los parámetros auxiliares² (β) del modelo son las magnitudes que nos permiten cuantificar ese influjo. En buena lógica debe ser el modelo la instancia en la que se especifique cuáles son las tareas observables que el modelo va a explicar, aunque en muchas ocasiones se construye para explicar tareas definidas con anterioridad, en función de criterios, a veces pragmáticos, a veces económicos, pero en cualquier caso exteriores a la teoría cognitiva referida al dominio.
- **Operacionalización.** Consiste en la asignación de valores a los parámetros auxiliares del modelo. En ocasiones esos valores se deducen de la propia teoría en la que se encuadra. Por ejemplo, en el modelo generativo de BEJAR para los tests de rotación mental que mencionamos más adelante, la dificultad está asociada a la disparidad angular entre las figuras. En ese caso los valores propuestos de los parámetros son una hipótesis derivada de la teoría. Una labor necesaria entonces consiste en el **contraste** de esa hipótesis. En otras ocasiones, las más numerosas, se desconoce el valor de los parámetros mencionados. Es necesario proceder por tanto a su **estimación**. El modelo está completamente operacionalizado cuando han sido determinados de uno u otro modo los valores de todos sus parámetros auxiliares, obteniéndose además en el proceso, información acerca de la incertidumbre asociada con esos valores.

2.1.1. Elementos de los modelos

Hay una estrecha relación entre el concepto de rasgo en la psicometría tradicional, las relaciones entre varios de ellos, y las técnicas empleadas para estudiar esas relaciones. Así, las estructuras de relaciones entre varios rasgos son siempre correlacionales, que es lo mismo que decir relaciones simétricas. En ese contexto no es posible hablar por ejemplo de relaciones de prerequisites o de contingencia.

El análisis cognitivo de los tests implica utilizar dos conceptos de gran utilidad. Se trata de las **tareas** cognitivas y las **habilidades**. Una tarea es una actividad que persigue un objetivo, tiene una duración temporal determinada, y existe un criterio por el que juzgar si se ha alcanzado el objetivo. Una tarea es observable. Una tarea puede ser simple o compleja. Es compleja cuando puede dividirse en varias sub-tareas, cada una con un objetivo parcial distinto de las demás. Las tareas pueden

² Ese conjunto de parámetros auxiliares serán, por ejemplo, los correspondientes a los items en la IRT. Su valor es estrictamente instrumental, puesto que el objetivo del diagnóstico es realizar inferencias sobre los sujetos, no sobre las tareas.

evaluarse como mínimo en dos estados, éxito o fracaso. Si una tarea compleja tiene varias subtareas, el estado de consecución de los objetivos de cada una de ellas determina los distintos estados que puede tomar la tarea compleja. Esos distintos patrones posibles determinan un conjunto ordenado, o al menos parcialmente ordenado.

Por su parte las habilidades son aquellos conocimientos o destrezas requeridos para llevar a cabo ciertas tareas específicas. La relación más simple consiste en una habilidad que es necesaria para realizar una tarea. Pero es posible que para realizar una tarea compleja sea preciso poseer varias habilidades, y también es posible que una habilidad permita realizar varias tareas. En el caso más sencillo una habilidad puede tener dos valores, presente o ausente. Una habilidad puede ser también compleja, y componerse de varias habilidades de orden inferior. Ese conjunto será como mínimo parcialmente ordenado. Pueden darse relaciones de sucesión, contingencia, contingencia conjunta, adyacencia, ligazón y no ordenamiento entre las habilidades, independiente de la estructura que tengan los patrones de las tareas (Haertel y Wiley, 1993):

Así como las tareas y sus grados de realización son manifiestas, las habilidades son latentes. El proceso de medida consiste por tanto en el establecimiento de una relación entre el patrón de estados manifiestos de las tareas, y el patrón de estados latente de las habilidades. En ese proceso, y desde este punto de vista cognitivo, es fundamental la fase de diseño del test, que no es ni más ni menos que la preparación de las tareas de forma que se haga patente la estructura latente de las habilidades. Haertel y Wiley (1993), proponen la creación de un álgebra que sirva para crear estructuras complejas de tareas que puedan ligarse a estructuras de habilidades. Este álgebra serviría para articular el papel de la estructura de tareas en la fase de asignación de puntuaciones, de forma que la los distintos estados o patrones de las habilidades se hiciesen manifiestos.

En definitiva un ítem es un conjunto de tareas. Esas tareas pueden identificarse con los ítems individuales, o puede que varias de ellas estén implicadas en un sólo ítem, o una sólo tarea implicar a varios ítems. Tanto las tareas como las habilidades latentes pueden ser de naturaleza dicotómica, o, pasando por los estados intermedios, de naturaleza continúa, en el otro extremo. Para cada una de las combinaciones que produce este conjunto de posibilidades, necesitamos un modelo psicométrico distinto. Algunos modelos existentes se han adaptado, y otros modelos nuevos han sido creados cuando ha sido preciso.

Una vez realizado el análisis cognitivo del dominio, el problema de la determinación de los patrones de respuesta que debe explicar el modelo se resuelve por medio del criterio de relevancia instructiva. Una vez determinadas las variables que sirven de criterio para la definición de los diferentes estados manifiestos, nos encontramos que la enumeración exhaustiva de las combinaciones de sus valores puede dar lugar a un universo de elementos de muy distinta relevancia instructiva. Hay dos fuentes de diferenciación. Por una parte tenemos los patrones producidos por errores sistemáticos en la aplicación de reglas o procedimientos. La otra fuente

son los errores aleatorios debidos posiblemente a factores espurios como falta de concentración, condiciones ambientales adversas, etc.

Una forma de resolver el problema del modelo de los diferentes estados cognitivos consiste en asignar a cada ítem un conjunto de atributos. Los estados se definen por las distintas combinaciones de presencia-ausencia de los atributos (Matriz Q): Aquí tenemos un problema, que es el de los falsos positivos y los falsos negativos. Un modelo determinístico no entraña estas consideraciones. Pero en un modelo probabilístico esto debe tenerse en cuenta. Entonces tenemos que a cada patrón de **tareas** no corresponde exáctamente un patrón de habilidades. Por eso debemos desarrollar algún método para detectar a partir de los patrones observados los patrones reales. Es decir, se trata de determinar qué patrones de tareas o respuestas, de entre las observadas, corresponden realmente a estados latentes de conocimiento.

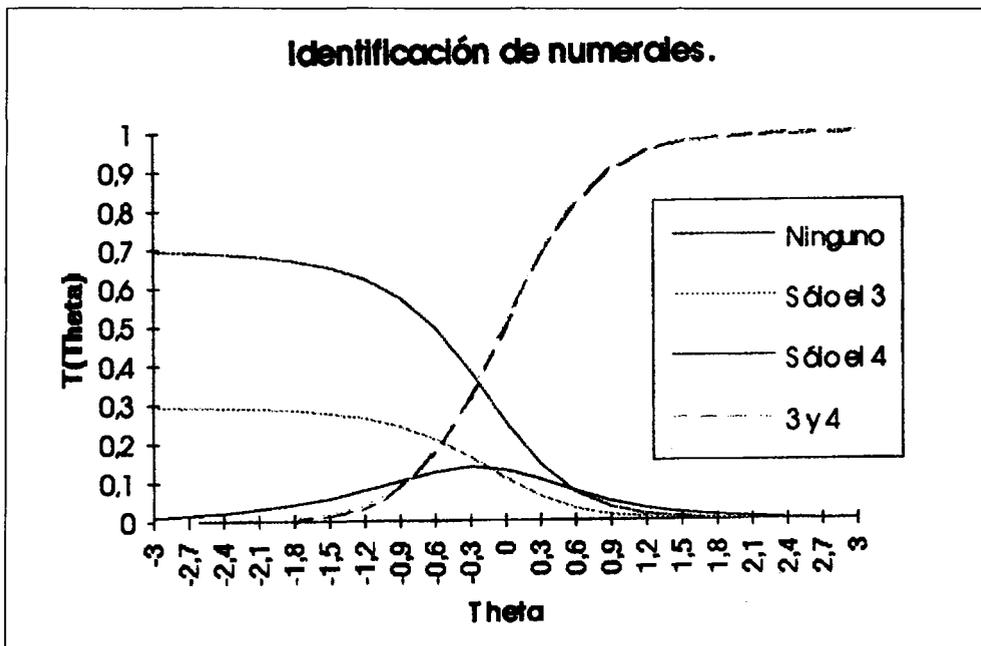
Dos soluciones distintas:

- Una ingenua, en la que el enrejado de patrones se determina a partir de los patrones más frecuentes. Se considera que los menos frecuentes se deben al error. (Haertel y Wiley, 1993).
- Otra más sofisticada, la de Tatsuoka (1990), consiste en asignar a cada patrón observado dos cantidades $\{\theta, \zeta\}$, siendo éste último un valor de «rareza». Es decir, dado un patrón de respuestas $\mathbf{x}=\{1,1,0,1,0,0,\dots,0,1\}$, a ese patrón le corresponde un valor de estimación máximo-verosímil de θ . Pero también se calcula un valor de que expresa lo raro que para el valor estimado de θ resulta el patrón de respuestas observado.

2.1.2. Estructura de los modelos

Cuando nos encontramos en el nuevo paradigma, algunos modelos desarrollados en el anterior pueden ser reinterpretados, adquiriendo un nuevo significado. Esto es algo muy común en la historia de la ciencia. Por ejemplo, la mecánica relativista tiene una capacidad explicativa mucho mayor que la mecánica Newtoniana. Pero en el nuevo marco establecido por la primera, la segunda se convierte en un caso particular, siendo de gran utilidad cuando los objetos mantienen velocidades pequeñas en relación a la velocidad de la luz. Cualquier moderno manual de física general presenta y desarrolla los principios de la mecánica clásica, si bien además incluirá la mecánica cuántica y la relativista.

Con relación a las nuevas ideas acerca del diagnóstico que estamos presentando, ocurre algo similar. Tenemos modelos que se han desarrollado en la fase anterior que resultan ser de utilidad en el nuevo paradigma. El «Multicomponent Latent Trait Model» resulta ser un ejemplo de modelo generativo, en el sentido señalado por Bejar (1993): El modelo explica y predice cómo se produce la dificultad del ítem en función de los componentes cognitivos que implica. Modelos en los que para su formulación ni siquiera se tuvo en cuenta la complejidad cognitiva del contexto pueden ser utilizados en una forma nueva. Los siguientes son dos ejemplos en esta línea.



En el primero, Thissen (1993), propone una recodificación de los datos, de forma que se crean pseudoítems, cuyas alternativas reflejan mejor algún constructo psicológico. Por ejemplo con la identificación y correspondencia de los numerales 3 y 4 con los conjuntos correspondientes (Bergan & Stone, 1985): A partir de varios ítems tradicionales crea dos pseudoítems. El primero se refiere a la identificación de los grafismos correspondientes al 3 y el 4. Las puntuaciones de ese pseudoítem son las categorías, «Ninguno», «Sólo el 3», «Sólo el 4», y «El 3 y el 4». El segundo se refiere a si los niños conocen la correspondencia entre los numerales y las cantidades de referencia. Las categorías del pseudoítem son las mismas. Para modelizar las respuestas a estos pseudoítems utiliza el modelo de Bock (1972) que ya ha sido mencionado con anterioridad.

$$T_x^*(\theta) = \frac{\exp(a_x\theta + C_x)}{\sum_{i=1}^m \exp(a_i\theta + C_i)}$$

Donde T_x es la curva de la opción x y m el número de opciones del ítem.

En otro ejemplo, (Klassen & o'Connor, 1987) sobre salud mental, se pretendía estudiar el comportamiento potencialmente violento de personas con historial de ese tipo de problemas. Se crean también dos pseudoítems. Uno con cuatro categorías referidas a la edad de la primera admisión por comportamiento violento, y otro con cuatro categorías referidas al número de eventos violentos anteriores. Para modelizar esto se usa el modelo de respuesta graduada de Samejima (1969): el modelo es

$T_x^*(\theta) = \frac{1}{1 + \exp(-a(\theta - b_{x-1}))}$ con las restricciones $T_1^* = 1$ y $T_{m+1}^* = 0$, siendo T_x^* la probabilidad de obtener la categoría x o superior. Las curvas divisorias para cada categoría se obtienen, tal y como hemos señalado con anterioridad, por

$$T_x(\theta) = T_x^*(\theta) - T_{x+1}^*(\theta).$$

(Para obtener la ecuación T_x^* , se codifican los datos con 1 si la respuesta fue x o superior, y con 0 si la respuesta fue inferior a x).

El único requisito para la creación de estos pseudoitems, es que las relaciones entre los datos observados estén explicadas por la variable latente (es decir, el supuesto de independencia local o condicional): En los dos casos se calcula las EAP(θ) («Expected a posterior»), que son las estimaciones de las puntuaciones θ correspondientes a cada combinación de valores.

Hay otros ejemplos en los que se desarrollan modelos psicométricos para reflejar una estructura cognitiva compleja. Por ejemplo el modelo híbrido (Yamamoto, 1987) supone la existencia de varios estados latentes discretos bien diferenciados entre sí, y una clase consistente en una dimensión continua para acomodar a los sujetos que no se adaptan bien a ninguna de las clases definidas. El dominio de conocimientos está por tanto dividido en dos subconjuntos. En el primero de ellos las clases están ordenadas. Las clases ordenadas están representadas por un modelo IRT. El otro subconjunto está formado por las clases no ordenadas, y están representadas por un modelo de clase latente en el que un conjunto de probabilidades condicionales caracterizan cada clase unívocamente, junto con una estimación de la proporción de individuos mejor descritos por cada una de las clases.

Supongamos que el modelo IRT adoptado es el de dos parámetros, siendo i el vector de parámetros (a_i, b_i) del ítem ζ_i . La probabilidad de respuesta correcta a dicho ítem dado θ_j , viene dado por el modelo logístico habitual

$$P(x_i = 1 | \theta_j, \zeta_i) = \frac{1}{1 + \exp(-Da_i(\theta_j - b_i))}$$

La probabilidad de una respuesta correcta al ítem i dada una clase latente k , se denota por $P(x_i=1|\gamma=k)$:

Si hacemos que la clase $\gamma=1$ corresponda a aquella modelizada por la función logística (IRT), entonces la probabilidad marginal del patrón de respuestas observadas x , bajo el supuesto de independencia condicional tanto en el modelo IRT como en el de clase latente, viene dada por

$$P(\mathbf{x} | \zeta) = \sum_{k=1}^K P(\mathbf{x} | \zeta, \gamma = k) P(\gamma = k)$$

$$= \int_{\theta} (\mathbf{x} | \theta, \zeta) f(\theta) d\theta P(\gamma = 1) + \sum_{k=2}^k P(\mathbf{x} | \zeta, \gamma = k) P(\gamma = k)$$

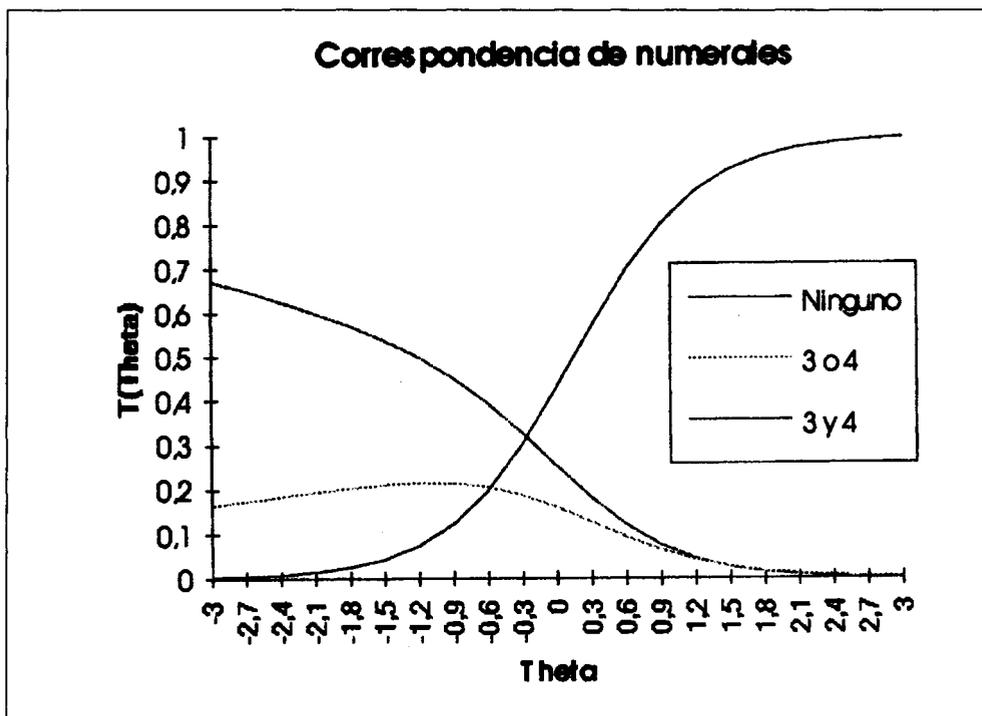


Figura 2
Segundo ejemplo de modelo de Bock para ítems politómicos.

Yamamoto (1991) realiza la evaluación de este modelo con los datos de una investigación de Gitomer y Van Slyke (1988), sobre la habilidad de técnicos electrónicos para interpretar símbolos de sistemas electrónicos digitales correspondientes a puertas lógicas.

El análisis cognitivo de las respuestas de los sujetos puso de manifiesto la existencia de errores sistemáticos en series de ítems que compartían ciertas características identificables. Estos errores sirvieron como base para la determinación de un conjunto de clases latentes. La pertenencia a determinada clase implica la reali-

zación sistemática de los mismos errores que los demás sujetos perteneciendo a esa misma clase. Otros sujetos no cometían errores sistemáticos que permitiese asignarlos a una de las clases previamente determinadas. El modelo híbrido descrito permite clasificar a los sujetos en uno de los grupos que cometían errores sistemáticos, o a un valor de la variable si eran asignados al grupo modelizado por IRT. Como resultado de esa clasificación es posible tomar decisiones instructivas para corregir las posibles concepciones erróneas de los sujetos acerca de las reglas de operación.

2.2. La construcción de los ítems

Es preciso diseñar situaciones de observación que permitan inferir el estado de comprensión de los sujetos. Serán los estímulos que provoquen las respuestas de los mismos, expresión, a su vez, de los niveles y naturaleza de su conceptualización en el dominio de que se trate.

El supuesto actuante es que los sujetos que difieren en estados de comprensión probablemente actúan de diferente modo, sus conductas observables son diferentes. Es fundamental, por tanto, la construcción de ítems que permitan poner de manifiesto estas diferencias. Serviría de poco el esfuerzo realizado en la especificación de modelos en la primera fase, si no se viera acompañado en esta segunda por la construcción de tareas que evidencien las variables implicadas.

La simple selección de los elementos de observación sólo en función de su contenido, no sería una actuación acorde con el objetivo propuesto. Que se conozcan sus demandas cognitivas es una exigencia razonable y coherente en el diseño del test o de las situaciones de observación en esta nueva orientación psicométrica. (Yamamoto, 1993; Snow y Lohman, 1993; Haertel y Wiley, 1993).

La base para inferir el estado de comprensión de un sujeto la proporcionan las probabilidades condicionadas de los diversos tipos de conducta que puede mostrar, dado el estado de comprensión que posea. En su expresión como $p(x|\eta)$, lo observado es x , y η los parámetros latentes.

Las observaciones x pueden ser de diferente naturaleza: evaluación de un experto, rapidez de una respuesta, características de cualquier realización del sujeto, etc.

Diferencias en estas probabilidades condicionadas, asociadas a diferentes configuraciones de los parámetros, reflejan la eficacia de cada ítem. Un ítem puede ser muy útil para distinguir entre algunos aspectos de los modelos potenciales de los sujetos pero menos para distinguir entre otros.

El tipo de ítem más oportuno depende de la inferencia a realizar.

No hay un solo «mejor» método para recopilar datos, sino sólo métodos más o menos efectivos. Tradicionalmente han sido los tests contruidos a base de ítems de opción múltiple los instrumentos más extensamente contruidos, utilizados y estudiados. Una reacción en contra de la estandarización es el movimiento de evaluación basado en las ejecuciones prácticas, conocido como «Performance Assessment». Como reacción pendular se pasa del extremo de la máxima igualación que permite la equiparabilidad de resultados, a la ausencia total de diseño previo de las

respuestas a las tareas propuestas. Realmente hay una cantidad ingente de concepciones distintas englobadas bajo este epígrafe. En cualquier caso es común a todas ellas el que no interesa tanto la solución al ítem como los procesos utilizados por el alumno para encontrarla. La respuesta que se pretende elicitar es una respuesta de construcción y elaboración más que de elección. Por ello los ítems aplicados son de ejecución práctica de tareas. El diseño de observación implica una distinta configuración del entorno de examen (Ruiz Primo, Baxter y Shavelson, 1993):

- una presentación de nuevos tipos de problemas.
- una provisión de materiales adicionales para realizar las posibles operaciones reclamadas (laboratorios, instrumentos de cálculo, etc.), al contrario que en los tests clásicos, que el sujeto se enfrentaba a las cuestiones con su bagaje personal de conocimientos.
- una realización de ejercicios que impliquen operaciones tales como búsqueda de fuentes, aplicación de técnicas aprendidas, realización de destrezas.

En este tipo de evaluación, una nota característica es que, al no haber un conjunto cerrado y predeterminado de respuestas posibles, no se puede establecer de antemano los criterios de valoración de las mismas. En teoría esas respuestas forman un conjunto infinito. Desde un punto de vista técnico la principal dificultad estriba en cumplir el requisito de equidad en la valoración de las distintas respuestas de los sujetos. El juicio que realizan el o los jueces debe encuadrarse en un marco conceptual común a todos ellos. De otra forma es imposible hablar de comparabilidad de resultados. Para lograr esta comunidad se hace imprescindible una seria y costosa labor de preparación de los evaluadores, a base de ensayos previos, de simulación, de análisis de casos, etc. Y eso en esencia es otra forma de estandarización.

Podemos concluir que la estandarización no es sinnónimo de opción múltiple. La estandarización tiene como objetivo lograr la igualación en las condiciones de evaluación, de forma que se disminuya la incertidumbre sobre los resultados. Pueden estandarizarse muchas facetas distintas; los ítems, las condiciones de aplicación, los criterios empleados por los jueces, etc. Las dificultades del «Performance Assessment» no vienen de la ausencia de estandarización, sino de la ausencia de un modelo cognitivo previo. El criterio que nos permite hablar de dos tipos distintos de evaluación diagnóstica es precisamente la existencia o no de un modelo cognitivo previo del alumno. Podemos hablar entonces mejor de evaluación con y sin modelo. Un buen ejemplo de cómo es posible la evaluación de ejecuciones prácticas, con un modelo previo, y la posibilidad de reducción de la incertidumbre en un contexto de «evaluación de dossiers» («portfolio assessment») de trabajos artísticos, lo constituye el trabajo citado por Mislevy (1993c) en el que se utilizan técnicas de control de calidad para detectar el «ruido» producido por las distintas fuentes actuantes en el proceso.

En cuanto a un formato más tradicional de evaluación, los ítems de opción múltiple, podemos decir que son muy eficientes en cuanto a coste y en la ordenación en una dimensión. Sin embargo es dudosa su utilidad en lo referido a la

identificación de estructuras de conocimiento y de los nodos o conexiones perdidas o erróneas. Se presentan en muchas ocasiones como alternativa las preguntas de respuesta abierta. Una comparación realizada por Birembaum y Tatsuoka (1987) entre ítems de opción múltiple y de formato abierto con resta de fracciones, demostró que la diferencia de formato no era importante si se juzgaba en términos tradicionales de fiabilidad y rendimiento medio general. Pero sí era muy importante si se juzgaba desde su diagnosticidad con relación a los errores de los estudiantes, a pesar de que los ítems de opción múltiple fueron contruidos para presentar los más comunes.

En general, las alternativas a los ítems de opción múltiple son mejores cuando existe **un modelo alternativo** de alumno y existe una relación lógica entre las tareas a observar, el modelo de alumno sobre el que inferir a partir de esas tareas, y el método de inferencia. No basta con cambiar de tarea a observar si no se cambia de modelo de alumno. Éste es un conjunto de estados de estructuras de conocimientos y destrezas cognitivas. Es una estructura formal. Y debemos ser conscientes de esto, porque si lo que hacemos es eliminar los ítems de opción múltiple o la exigencia de mantener ciertas características técnicas de nuestros instrumentos de medida para seguir puntuando a los alumnos en una escala de cero a diez, no hemos cambiado de modelo latente de alumno, pero hemos rebajado la fiabilidad y la validez de nuestro proceso de medida. Y parafraseando a Messick (1992), no debemos olvidar que la validez no es sólo un problema de relevancia técnica, sino de relevancia social. El que un instrumento sea fiable y válido es un problema de naturaleza moral, de responsabilidad ética de quien lo construye y aplica, en definitiva de quien hace el diagnóstico, individualmente, con cada uno de los alumnos que lo responden, y en conjunto, con la sociedad a la que se sirve.

Por todo ello la naturaleza de las tareas que el tests o el diseño de la situación de observación debe incluir vendrá determinada por las necesidades inferenciales del proceso diagnóstico. Así, mientras que las teorías estándar de los tests se centran en las respuestas a los ítems, no en su contenido, la psicología cognitiva exige el centrar la atención sobre el contenido de los ítems, y su interrelación con las respuestas. El análisis cognitivo del contenido y la determinación de tareas y destrezas implicadas en los ítems, es una característica de esta tendencia.

Por otra parte, en el enfoque tradicional un buen ítem es aquél que discrimina bien entre sujetos. Ahora es preciso que un ítem aporte «diagnosticidad», es decir, capacidad diagnóstica. Es preciso saber por qué un alumno responde mal a un ítem dado, qué proceso cognitivo ha fallado, qué tarea no se ha realizado, qué conocimiento faltaba, o qué componente cognitivo no estaba presente. Esto hace que en los nuevos modelos haya una mayor focalización en los tests como conjuntos jerarquizados, interrelacionados de ítems. La posibilidad de realizar diagnóstico cognitivo depende de la información proporcionada por varios ítems simultáneamente. Y esos ítems no pueden provenir de un pool «amorfo» y unidimensional. El test se convierte en una estructura de tareas, en una red de ítems en la que «pescar» los distintos estados de los alumnos.

En un intento de proponer un marco teórico general que sirva para construir los ítems que deberán ser incluidos en un determinado instrumento diagnóstico, Bejar (1993) propone lo que denomina «Response Generative Modeling», implicando

(...) una gramática, capaz de asignar una descripción psicométrica a cada ítem en un universo de ítems, y también capaz de generar todos los ítems en ese universo.

Esto implica:

1. Más fácil interpretación de las puntuaciones, ya que sabemos cómo ha sido generado el ítem.
2. Un conocimiento del proceso completo de respuesta al ítem. (Esto permitiría abandonar el formato de opción múltiple).
3. Capacidad de desarrollar ayudas inteligentes a la generación de tests.

La idea central es, no sólo que se establezcan las reglas por las cuales se generan los ítems, sino que además estos ítems se generan con las características psicométricas, tal como la dificultad, deseadas.

Esto significa que la postulación de procesos y estructuras de conocimientos subyacentes requeridas para responder a un ítem es no sólo admisible, sino el corazón mismo de este enfoque. (Bejar, 1993, p. 326).

Los antecedentes de este tipo de modelización hay que buscarlos en el «Item Sampling Modeling» (Tryon, 1957): Estaba basada más en la competencia que en el rendimiento, en el sentido en que Chomsky (1965) diferenciaba entre **competencia** y **rendimiento** lingüísticos. Se daba la paradoja de que algunos ítems que estaban diseñados para medir un mismo objetivo educativo (Competencia), tenían distinta dificultad y distinta proporción de sujetos respondiéndolos correctamente. Eso naturalmente no tiene explicación en ausencia de un componente de rendimiento.

Una explicación de la dificultad de los ítems aumentaría la validez del test, ya que tendría que incorporar la estructura mental y los procesos necesarios para responder al test. Si ese modelo ha sido derivado de una teoría que tiene apoyo empírico, entonces el estatus de validez de las puntuaciones de ese test es superior a otro test en el que las características de los ítems no se conocen hasta que se aplica el test.

La validez es considerada en la concepción tradicional como explicación de la consistencia de las respuestas. Se trata de determinar qué proporción de la varianza es explicada por el rasgo medido, y qué proporción lo es por el método común empleado para medir el rasgo. (Es decir, qué proporción de varianza se debe a variables sustantivas y cuál a variables irrelevantes, como el método de medida (Campbell y Fiske, 1959)).

En la nueva concepción la validez es considerada, además, como la explicación de la dificultad de los ítems. En qué proporción la dificultad se debe a factores espurios y en qué proporción se debe a la complejidad cognitiva. Esto implica, no nos cansaremos de resaltarlo, la introducción de un elemento causal en los modelos psicométricos.

La aproximación generativa a la construcción de ítems supone dos cosas importantes: a) la existencia de un mecanismo de generación de los ítems, y b) un conocimiento suficiente acerca del proceso de respuesta a los mismos que permita estimar los parámetros psicométricos de los ítems generados.

Hay varios ejemplos de la posibilidad de este enfoque. Posiblemente el más completo se refiere a la habilidad espacial. En concreto la rotación mental de figuras geométricas es un campo ampliamente estudiado. Está bien establecido (Corballis, 1982) que la disparidad angular entre dos figuras rotadas, determina en gran medida el tiempo de respuesta. El control de la dificultad de un ítem de este tipo por tanto se lograría a través de la disparidad angular entre las dos figuras propuestas. En un test adaptativo de este tipo, a cada sujeto se le presentarían dos figuras con mayor o menor ángulo de separación en función de la respuesta a los ítems anteriores. El ordenador encargado de presentar los estímulos mediría el tiempo que el sujeto tarda en dar su respuesta. A partir del tiempo que el sujeto tarda en proporcionar la respuesta correcta, podemos estimar la dificultad relativa que representa ese ítem para el sujeto. Gaviria (en preparación) propone un modelo muy simple para el tiempo que se tarda en dar una respuesta correcta.

$$t_{ij} = \tau_{ij} + \mu_i \exp(-a(\theta_j - b_i)) + e_{ij}$$

donde t_{ij} es el tiempo que tarda el sujeto j en dar la respuesta correcta al ítem i , τ_{ij} es el tiempo que se tarda en responder correctamente al ítem i cuando θ_j es infinitamente mayor que b_i . Es decir, es en la práctica el tiempo de procesamiento y respuesta mecánica mínimo necesario. El parámetro μ_i es un factor de escala. Es el tiempo que se tarda en responder correctamente cuando $\theta_j - b_i = 1$. Por último e_{ij} es una variable aleatoria, con distribución log-normal.

Hay otros dominios en los que no existe un grado de conocimiento suficiente de los procesos psicológicos que subyacen al proceso de respuesta a un tipo determinado de ítems. En ese caso podrían relajarse un poco las exigencias impuestas al proceso de construcción, y en vez de generar los valores específicos de la dificultad, bastaría con establecer un orden entre las dificultades de los distintos ítems construidos.

Bejar (1993), presenta ejemplos de varios ámbitos distintos y en distinto grado de realización en los que puede apreciarse las posibilidades de este concepto.

Duncan Y Humphreys (1989), proponen una teoría que explica el rendimiento en tareas de localización de figuras escondidas.

Butterfield, Nielse, Tangen y Richardson (1985), aportan evidencia empírica de gran calidad sobre la posibilidad de un enfoque generativo en los test de razonamiento inductivo basados en series de letras.

Johnson-Laird, Byrne, y Tabossi (1989), ilustran las posibilidades de este mismo enfoque en tests de razonamiento deductivo.

Las posibilidades de una modelización generativa en el área del razonamiento analógico son estudiadas por Bejar, Chaffin y Embretson (1991): La generación de analogías ha sido demostrada por Chaffin y Hermann (1987):

Roid y Haladyna (1982) han trabajado sobre los mecanismos de generación de ítems en tests de razonamiento aritmético y cuantitativo. En este mismo campo se han estudiado las variables que afectan a la dificultad de los ítems. El resultado de esas investigaciones puede encontrarse en el volumen editado por Goldin y McClintock (1984): La generación automática de ítems en este dominio ha sido tratada por Hively, Paterson, y Page (1968) y Brown y Burton (1978): En el mismo terreno del razonamiento aritmético y cuantitativo pueden encuadrarse los trabajos de Jarjoura y Brennan (1982) y Kolen y Harris (1987):

En el dominio verbal, Bejar (1988) trata de un sistema para evaluar automáticamente la capacidad escritora. El sistema se basa en un motor de corrección gramatical.

Katz (1988), en el terreno de la comprensión lectora, ha desarrollado un sistema (START) que analiza automáticamente cualquier texto en inglés, y lo transforma en una representación proposicional de tal forma que pueden generarse fácilmente preguntas sobre un texto.

Brown y Burton (1978), ya en el ámbito de los tests de rendimiento, hacen un análisis del trabajo con fracciones resaltando la importancia de los errores sistemáticos y su determinación a partir de las repuestas a ciertos ítems generados para detectarlos.

Otros trabajos se centran en la consideración de un enfoque generativo en destrezas de detección de averías en sistemas electrónicos, o en el diagnóstico de enfermedades. Para este tipo de destrezas véase Kieras (1990) Fulton y Pepe (1990), Lesgold, Ivill-Friel, y Bonar (1989), Parker y Miller (1988) Pearl (1987), Miller (1984) Braun, Carlson y Bejar (1989) y Warner y asociados (1988).

2.3. La inferencia

La inferencia consiste en enlazar las observaciones x con los estados hipotetizados en el modelo. Se logra de ese modo extraer las conclusiones pertinentes sobre la naturaleza y niveles de comprensión de los alumnos, conclusiones que son la base para la toma de decisiones educativas, y se cierra así el proceso que caracteriza cualquier aplicación diagnóstica en esta nueva orientación.

Tenemos dos niveles distintos de inferencia en diagnóstico. Hemos determinado que el proceso de construcción de un modelo diagnóstico está completo cuando el modelo está operacionalizado, es decir, incluye las herramientas para poner en relación lo observado con lo latente. Nos referimos a las probabilidades condicionales.

Un primer nivel de inferencia se produce en este punto. A partir de los datos observados debemos «estimar» los valores de los parámetros auxiliares del modelo.

Estos parámetros nos permiten establecer la probabilidad de cada patrón de conductas o respuestas observadas, condicionada a los estados latentes del sujeto. Esta fase es en realidad un paso más, aunque muy importante, en el proceso de operacionalización del modelo.

El otro nivel de inferencia es el que define esencialmente la naturaleza del proceso diagnóstico. Consiste en determinar la probabilidad de cada estado latente en el espacio en el que se encuadra un sujeto a partir de la evidencia empírica que obtenemos en el proceso de evaluación. Se trata de la determinación de $P(\eta|\mathbf{x})$. La revisión de las probabilidades se realiza mediante $P(\eta|\mathbf{x}) \propto P(\mathbf{x}|\eta)P(\eta)$: El término $P(\mathbf{x}|\eta)$ viene dado por el modelo psicométrico formulado. $P(\eta)$, distribución a priori de η , refleja nuestro conocimiento acerca de los estados latentes. En ocasiones puede tratarse de una distribución no informativa, por lo que la estimación será equivalente a la de máxima verosimilitud. Calcular el término de la izquierda es la esencia del proceso de diagnóstico. En este proceso, a partir de los datos observados, y por medio de la inferencia que nos permite llevar a cabo nuestro modelo, afirmamos la hipótesis de que el alumno presenta determinada configuración cognitiva. Esa hipótesis es mantenida hasta que obtenemos cierta evidencia que la niega o que la modifica en parte.

Como señala Messick (1989), la validez hace referencia precisamente a la fuerza con la que pueden llevarse a cabo las inferencias que acabamos de mencionar. Tiene que ver por tanto con el modelo psicológico que ha servido para definir qué tareas eran la evidencia empírica necesaria para realizar la inferencia diagnóstica. No tiene que ver con el instrumento en sí, sino con la relación entre la naturaleza del instrumento y la teoría que lo determina.

2.3.1. *Redes inferenciales*

Los modelos cognitivos que han de servir de base para el proceso de diagnóstico, distan mucho de la simplicidad formal que subyace en la teoría estándar de los tests. En los modelos cognitivos tenemos una gran cantidad de variables y estados relacionados entre sí. Realizar inferencias en esos sistemas podría ser una tarea imposible de ser llevada a cabo en tiempo real. En los últimos años el estudio de estos sistemas ha florecido debido a su utilidad en la implementación de sistemas expertos. Las redes inferenciales (Lauritzen & Spiegelhalter, 1988; Pearl, 1988) se basan en la idea de utilizar las relaciones de dependencia condicional entre las variables. Los primeros desarrollos se han centrado en sistemas expertos en el diagnóstico médico, como MUNIN. Posteriormente el «motor inferencial» del sistema ha sido aislado para poder ser aplicado a otros campos.

Una red inferencial es un sistema probabilístico, en esencia, es un modelo causal. En él, los valores de ciertas variables observables están condicionados a los valores de otras variables no observables que son las que explican a las primeras. La topología de la red es tal que se asegure la existencia de independencia condicional entre las variables observadas. La dirección causal se postula desde las variables no

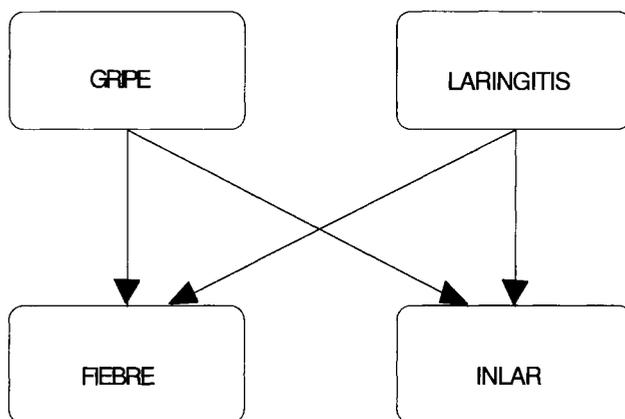


Figura 3
Red causal de dos enfermedades y dos síntomas.

observadas a las observadas. La probabilidad de cada estado observado del sistema está condicionado por el estado de las variables latentes. Sin embargo, es el estado latente el que nos interesa diagnosticar, y para ello debemos inferirlo a partir de las observaciones realizadas. Estos cálculos inferenciales se han realizado tradicionalmente en sistemas causales similares por medio de estimaciones de máxima verosimilitud. Recientemente, (Lauritzen y Spiegelhalter, 1988; Pearl, 1988; Shafer y Shenoy, 1988), se han desarrollado algunas generalizaciones del teorema de Bayes que permiten llevar a cabo las actualizaciones de las probabilidades asociadas a cada estado del sistema a base de varios conjuntos independientes de actualizaciones locales de dichas probabilidades. Desarrollamos brevemente el ejemplo presentado por Mitlevy (1993b) para ilustrar el método.

Supongamos que tenemos cuatro variables, dos de ellas actuando como causas, y otras dos como efectos. En el ejemplo que nos ocupa, las dos causas son dos enfermedades, «gripe» y «laringitis», y los dos efectos son dos síntomas asociados a ambas enfermedades, la fiebre y la inflamación de laringe. Por simplicidad asumimos, sin pérdida de generalidad, que estas cuatro variables sólo tienen dos valores, «presencia» o «ausencia».

En la figura 3 aparece un diagrama causal de esta red. En la práctica diagnóstica ordinaria, nos encontramos con los síntomas, presencia o ausencia de fiebre, y presencia o ausencia de inflamación de laringe, y a partir de ellos nos interesa determinar cuál es la probabilidad asociada a los valores de presencia o ausencia de cada una de las enfermedades posibles causas de esos síntomas. La observación de uno solo de los síntomas modifica las probabilidades asociadas a cada una de las causas, y esto a su vez influye en nuestras expectativas respecto a la aparición del otro síntoma. Naturalmente, en sistemas reales esto representa una complejidad potencial muy grande, necesitando por lo tanto algoritmos de actualización sumamente eficaces, para que la evaluación de la situación pueda hacerse en tiempo real.

El procedimiento desarrollado por Lauritzen y Spiegelhalter, se basa en la identificación de unos subconjuntos de variables denominados «cliques».

Las siete fases son:

1. Representación recursiva de la distribución conjunta de todas las variables implicadas.
2. Representación de (1) en un grafo direccional.
3. Grafo triangulado no direccionado.
4. determinación de los cliques y sus intersecciones.
5. Representación del árbol conjunto.
6. Tablas potenciales.
7. Esquema de actualización.

La representación recursiva de la distribución conjunta de un grupo de variables aleatorias, $X_1 \dots X_n$, toma la forma³,

$$P(X_1, \dots, X_n) = P(X_n | X_{n-1}, \dots, X_1) P(X_{n-1} | X_{n-2}, \dots, X_1) \dots P(X_2 | X_1) P(X_1) = \prod_{j=1}^n P(X_{n-j+1} | X_{n-j}, \dots, X_1)$$

En el caso que nos ocupa, tendríamos

$$P(\text{fiebre, inlar, laringitis, gripe}) = P(\text{fiebre} | \text{inlar, gripe, laringitis}) P(\text{inlar} | \text{gripe, laringitis}) P(\text{gripe} | \text{laringitis}) P(\text{laringitis})$$

Dado que las variables latentes, laringitis y gripe, inducen independencia condicional entre las variables observadas, tenemos que $P(\text{fiebre} | \text{inlar, laringitis, gripe}) = P(\text{fiebre} | \text{laringitis, gripe})$: Por otra parte en el modelo se asume que laringitis y gripe son independientes entre sí. Si no lo fueran, bastaría con incluir en el modelo una variable auxiliar que indujese independencia condicional entre ellas. Dada esa independencia también tenemos que $P(\text{gripe} | \text{laringitis}) = P(\text{gripe})$: Con estas equivalencias la distribución conjunta queda, $P(\text{fiebre, inlar, laringitis, gripe}) =$

$$= P(\text{fiebre} | \text{gripe, laringitis}) P(\text{inlar} | \text{gripe, laringitis}) P(\text{gripe}) P(\text{laringitis}).$$

La reconstrucción de la distribución conjunta implica la determinación de los cuatro elementos de esta expresión. Por una parte tenemos $P(\text{gripe})$ y $P(\text{laringitis})$: Estas probabilidades son las distribuciones a priori de las enfermedades, que por lo general serán equivalentes a la proporción de la población que en un determinado momento sufre cada una de ellas. Sin embargo los valores de $P(\text{fiebre} | \text{gripe, laringitis})$ y $P(\text{inlar} | \text{gripe, laringitis})$, provienen de dos fuentes alternativas. O bien se deducen

³ Esta expresión se obtiene fácilmente de $P(X_1, X_2) = P(X_1 | X_2) P(X_2)$, sustituyendo X_2 por (X_2, X_3, \dots, X_n) .

de una teoría previa acerca de las relaciones entre las enfermedades y cada uno de los síntomas, o bien los conocemos empíricamente a partir de experiencia anterior.

En las tablas 1 y 2 tenemos, respectivamente, las probabilidades a priori de las enfermedades y las probabilidades condicionales para los síntomas del ejemplo que nos ocupa.

TABLA 1
PROBABILIDADES A PRIORI DE LAS ENFERMEDADES

	GRUPE	LARINGITIS
SÍ	0,11	0,11
NO	0,89	0,89

TABLA 2
PROBABILIDADES DE LOS SÍNTOMAS CONDICIONADOS POR LAS ENFERMEDADES

Gripe	Laringitis	P(Inlar=SÍ)	P(Inlar=NO)
SÍ	SÍ	0,91	0,09
SÍ	NO	0,05	0,95
NO	SÍ	0,90	0,10
NO	NO	0,01	0,99

Gripe	Laringitis	P(Fiebre=SÍ)	P(Fiebre=NO)
SÍ	SÍ	0,99	0,01
SÍ	NO	0,90	0,10
NO	SÍ	0,90	0,10
NO	NO	0,01	0,99

Las fases 1 y 2 están estrechamente relacionadas. De hecho la segunda es la expresión gráfica de la primera. La fase 3 tiene como objeto la determinación de los cliques del grafo y de las intersecciones entre cliques. Un clique es un subconjunto de variables que están ligadas dos a dos, y cuya función consiste en evitar las complicaciones de cálculo que supone la existencia de lazos de retroalimentación. En el ejemplo que estamos analizando, se dan también este tipo de lazos. Por ejemplo, conocer el estado de fiebre de un paciente modifica las probabilidades de los estados de gripe, que a su vez modifica las probabilidades de la inflamación de laringe, que a su vez modifica las probabilidades de laringitis, que afecta simultáneamente a las probabilidades de fiebre. En sistemas complejos estos lazos complican extraordinariamente el proceso de actualización de probabilidades. Como vere-

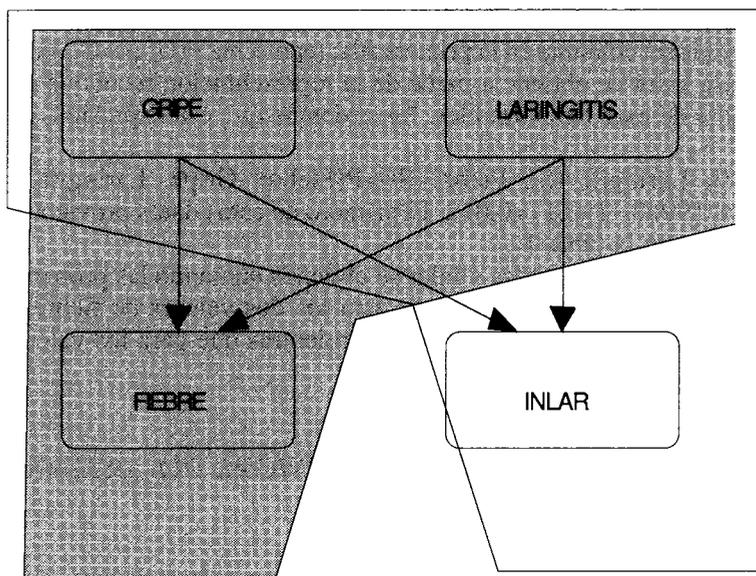


Figura 4
Cliques del grafo.

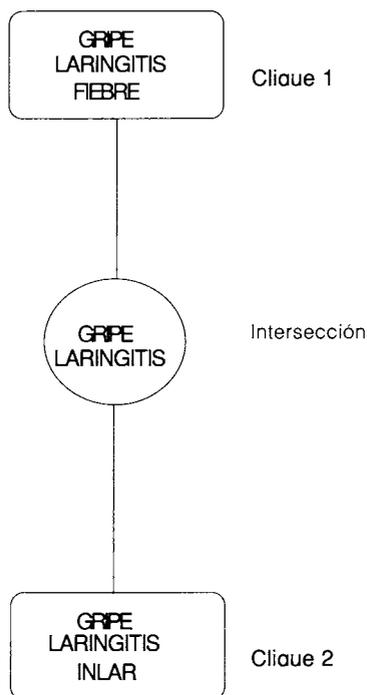


Figura 5
Árbol de representación conjunta.

mos, la utilización de los cliques proporciona un método eficiente de propagación de la información por todo el sistema.

En el presente ejemplo hay dos cliques, {Fiebre, Gripe, Laringitis} y {Gripe, Laringitis, Inlar}, y una intersección de cliques, {Gripe, Laringitis}. Como vemos, en la intersección de cliques se incluyen todas las variables que están incluidas simultáneamente en dos cliques distintos. En la figura 2 vemos identificados los cliques sobre el grafo.

La figura 5 representa el árbol conjunto. Este grafo es la herramienta que dirige el proceso de actualización. En él aparecen representados cada uno de los cliques y su intersección.

Las tablas potenciales son unas tablas de probabilidades condicionales que recogen la estructura del árbol conjunto, y constituyen el medio de actualización local. Estas tablas reflejan la situación del sistema en cada momento, y van propagando los cambios que se producen desde un clique a los demás a través de las intersecciones.

En la tabla 3 tenemos las probabilidades iniciales de nuestro caso. Por ejemplo, la casilla de la tercera columna en la primera fila, representa $P(\text{Gripe}=\text{Sí}, \text{Laringitis}=\text{Sí}, \text{Fiebre}=\text{Sí})$: Ese valor se obtiene a partir de la representación recursiva de la distribución conjunta de esas tres variables. Recordemos que esa expresión es

$P(\text{Gripe}=\text{Sí}, \text{Laringitis}=\text{Sí}, \text{Fiebre}=\text{Sí})=P(\text{Fiebre}|\text{Gripe}, \text{Laringitis}) P(\text{Gripe}|\text{Laringitis})=0,99 \times 0,11 \times 0,11=0,012$. El primero de estos datos proviene de la tabla 2, y los otros dos de la tabla 1.

Los valores iniciales de Fiebre e Inlar se obtienen sumando las probabilidades de su columna correspondiente. Eso nos da para los dos valores de fiebre (Sí y No), unas probabilidades iniciales de 0,20 y 0,80, mientras que para los valores de Inlar obtenemos 0,112 y 0,888.

TABLA 3
TABLA POTENCIAL DEL ESTADO INICIAL DEL SISTEMA

Clique 1

Gripe	Laringitis	Fiebre=Sí	Fiebre=No
Sí	Sí	0,012	0,000
Sí	No	0,088	0,010
No	Sí	0,088	0,010
No	No	0,008	0,784
Intersección			
Gripe	Laringitis	Probabilidad	
Sí	Sí	0,012	
Sí	No	0,098	
No	Sí	0,098	
No	No	0,792	
Clique 2			
Gripe	Laringitis	Inlar=Sí	Inlar=No
Sí	Sí	0,011	0,001
Sí	No	0,005	0,093
No	Sí	0,088	0,010
No	No	0,008	0,784

Veamos ahora cómo se produce una actualización. Supongamos que obtenemos el dato de que un paciente tiene fiebre. Automáticamente la primera columna del primer clique tiene que sumar uno, y todos los valores de las casillas de la segunda columna se ponen a cero. En realidad para el proceso de propagación no es necesario normalizar a 1 las probabilidades de la primera columna, ya que lo que se propaga es la proporcionalidad entre las casillas.

A continuación las probabilidades de la primera columna se propagan a la tabla de la intersección. Desde esa tabla se procede a actualizar las probabilidades de la tabla del segundo clique. Para ello se procede como sigue. En cada fila de la tabla del clique 2 dividimos los valores de cada una de las casillas por el valor que tenía la casilla correspondiente del clique de la intersección antes de producirse la primera propagación y multiplicándolo por el nuevo valor de esa misma casilla. En este caso, por ejemplo, en la tercera casilla del clique 2 dividimos 0,088 por 0,098, obteniendo 0,898. Ese valor se multiplica por 0,088, obteniéndose 0,079. (En la tabla se ha redondeado al tercer decimal):

El último paso consiste en normalizar las probabilidades de las casillas del clique 2 para la suma de los totales de las dos columnas sea 1.

La lógica del proceso presentado es la siguiente. Llamemos $P_0(\text{Gripe, Laringitis, Inlar})$ a las probabilidades del clique 2 al inicio del proceso. Estas probabilidades son consistentes con las correspondientes del clique 1 en el mismo momento. El nuevo dato acerca de la fiebre nos hace definir $P_1(\text{Gripe, Laringitis})$ con los nuevos valores. Si llamamos $P_1(\text{Gripe, Laringitis, Inlar})$ a la distribución conjunta del segundo clique después de la propagación, se cumplirá que

$$P_1(\text{Gripe, Laringitis, Inlar}) = P(\text{Inlar}|\text{Gripe, Laringitis})P_1(\text{Gripe, Laringitis})$$

Como queremos poner $P_1(\text{Gripe, Laringitis, Inlar})$ en función de $P_0(\text{Gripe, Laringitis, Inlar})$ y de $P_1(\text{Gripe, Laringitis})$, podemos hacerlo sabiendo que

$P_0(\text{Gripe, Laringitis, Inlar}) = P(\text{Inlar}|\text{Gripe, Laringitis})P_0(\text{Gripe, Laringitis})$ luego $P(\text{Inlar}|\text{Gripe, Laringitis}) = P_0(\text{Gripe, Laringitis, Inlar})/P_0(\text{Gripe, Laringitis})$: Sustituimos en la primera expresión, y

$$\begin{aligned} P_1(\text{Gripe, Laringitis, Inlar}) &= \\ &= \{P_0(\text{Gripe, Laringitis, Inlar})/P_0(\text{Gripe, Laringitis})\}P_1(\text{Gripe, Laringitis}): \end{aligned}$$

Una explicación detallada del proceso completo puede encontrarse en Mislevy (1993b) y Lauritzen y Spiegelhalter, (1988):

TABLA 4
TABLA POTENCIAL DEL SISTEMA PARA FIEBRE=SÍ

Clique 1

Gripe	Laringitis	Fiebre=Sí	Fiebre=No
Sí	Sí	0,012	0,000
Sí	No	0,088	0,000
No	Sí	0,088	0,000
No	No	0,008	0,000
Intersección			
Gripe	Laringitis	Probabilidad	
Sí	Sí	0,012	
Sí	No	0,088	
No	Sí	0,088	
No	No	0,008	
Clique 2			
Gripe	Laringitis	Inlar=Sí	Inlar=No
Sí	Sí	0,011	0,001
Sí	No	0,004	0,084
No	Sí	0,080	0,009
No	No	0,000	0,008
Clique 2			
Gripe	Laringitis	Inlar=Sí	Inlar=No
Sí	Sí	0,059	0,005
Sí	No	0,020	0,426
No	Sí	0,406	0,046
No	No	0,000	0,041

2.4. Un ejemplo de diagnóstico educativo

Podemos ver integradas las distintas fases del proceso de diagnóstico, en un ejemplo muy interesante (Mislevy, 1991) referido a ciertas habilidades implicadas

en un experimento cognitivo diseñado por Siegler (1981), para evidenciar los diferentes estados de comprensión. Las tareas planteadas por Siegler están referidas a un clásico experimento piagetiano en el que se presenta a los niños una balanza de brazos iguales afirmados por dos topes, y después de poner diversas combinaciones de pesos y distancias se les pide que predigan el movimiento de los brazos cuando se retiren los topes.

Fruto de las investigaciones de Piaget, tenemos una definición de posibles estados de comprensión de los elementos actuantes en el experimento y la secuencia habitual de esos estados. Tal conocimiento ayuda a precisar los componentes del modelo cognitivo, es decir, las variables latentes que deseamos estudiar. Es precisamente el modelo cognitivo el que orienta, además, el proceso de definición de las tareas si, como se pretende, han de poner de manifiesto esos diferentes estados.

La progresión habitual de la comprensión de los niños sobre el funcionamiento de la balanza se define en función de la adquisición de ciertas reglas que combinan la relación entre la magnitud de los pesos ubicados a ambos lados de la balanza y sus distancias respecto del «fiel». Estas reglas son las siguientes:

Regla I: Si son iguales los pesos a los dos lados del fiel, la balanza permanecerá en equilibrio. Puesto que los niños comprenden la importancia del peso para la resolución del problema antes que la distancia, la primera se conoce como «dimensión dominante» y esta última como «dimensión subordinada».

Regla II: Si los pesos y distancias son iguales en los dos lados, la balanza permanecerá en equilibrio. Si los pesos son iguales, pero no las distancias, se inclinará hacia el lado con mayor distancia. Del mismo modo, el lado que tenga mayor peso hará que la balanza se incline hacia él. La aplicación de esta regla implica que el niño utiliza fundamentalmente la información de la dimensión dominante, y sólo cuando no pueda decidir basándose en ella, utilizará la información de la dimensión subordinada.

Regla III: Si los valores de los pesos y de las longitudes son desiguales en los dos lados, el niño saldrá del paso respondiendo al azar; y ello porque si bien comprende que las dos dimensiones son importantes, no sabe cómo combinarlas.

Regla IV: Combina pesos y distancias correctamente.

Se hipotetiza, por tanto, que cada niño se puede clasificar en alguno de estos cuatro estados de comprensión o en un quinto, denominado «preoperacional», en el que el chico no reconoce que la acción de la balanza tenga que ver con los pesos y las distancias.

Siegler diseñó seis tipos de ítems que recogían las tareas en las que se debía reflejar el nivel alcanzado en la comprensión de estas reglas. Las variaciones en la magnitud de los pesos colocados a cada lado de la balanza y las distancias a las que éstos se ubican a partir del «fiel», se hicieron de forma que reflejasen distintas relaciones entre las magnitudes denominadas dominante y subordinada. Estas relaciones se clasificaron en las siguientes categorías.

Igualdad (E): La información de las dos dimensiones por separado es coherente, e indica que la balanza permanecerá en equilibrio.

Dominancia (D): La distancia es la misma, pero la información de la dimensión dominante indica que la balanza se desequilibrará.

Subordinación (S): Los pesos son iguales pero no las distancias, con lo que el uso de la información de la dimensión subordinada permite determinar que la balanza se desequilibrará hacia el lado de mayor distancia.

Conflicto-dominancia (CD): La información de las dos dimensiones por separado es contradictoria, y el uso de la información de la dimensión dominante permite predecir correctamente que la balanza se desequilibrará hacia el lado de mayor peso.

Conflicto-subordinación (CS): La información de las dos dimensiones es contradictoria, y el uso de la dimensión subordinada permite predecir que la balanza se desequilibrará hacia el lado de mayor distancia.

Conflicto-igualdad (CE): La información de las dos dimensiones por separado es contradictoria, pero el uso correcto de las relaciones entre ambas permite predecir que la balanza permanecerá en equilibrio.

Esto dio lugar a los seis tipos de ítems: E, D, S, CD, CS, CE, de los que pueden verse un ejemplo de cada clase en la figura 6.

Obtenidos el modelo de alumno y las tareas observables, necesitamos determinar cuál es la relación que existe entre esos elementos, de forma que a partir de las observaciones podemos hacer inferencias acerca de los estados latentes. Si para ello utilizásemos el modelo cognitivo directamente, que es un modelo determinístico, tendríamos que suponer que las respuestas de los sujetos están en perfecta concordancia con las reglas hipotetizadas. En la tabla 5 vemos las probabilidades condicionales para ese modelo.

TABLA 5
PROBABILIDADES CONDICIONALES TEÓRICAS

Tipo de problema	Estado 0	Estado I	Estado II	Estado III	Estado IV
E	0,333	1,000	1,000	1,000	1,000
D	0,333	1,000	1,000	1,000	1,000
S	0,333	0,000	1,000	1,000	1,000
CD	0,333	1,000	1,000	0,333	1,000
CS	0,333	0,000	0,000	0,333	1,000
CE	0,333	0,000	0,000	0,333	1,000

La observación de esta tabla nos permite el siguiente comentario. Por ejemplo, una respuesta correcta a un ítem del tipo S nos llevaría a concluir que el estado de comprensión del sujeto correspondiente no es el I, puesto que en ese estado la probabilidad de respuesta correcta a ítems de ese tipo es nula.

Sin embargo, las inconsistencias que a menudo caracterizan las respuestas de los sujetos, pueden provocar algunas observaciones inesperadas, tanto en un sentido

como en otro: falsos positivos y falsos negativos. Los primeros son contestaciones correctas dadas por sujetos que no poseen los requisitos exigidos de comprensión. Contestar bien por razones falsas. El sujeto realiza adivinaciones afortunadas. Los falsos negativos son errores en un ítem cometido por sujetos que tienen los requisitos necesarios de comprensión.

Además, los sujetos pueden tener éxito en unas tareas y fracasar en otras que requieran las mismas habilidades. Esta variedad de posibilidades nos lleva, en resumen, a esperar cualquier tipo de respuesta, aunque con distintos niveles de probabilidad, de chicos en diferentes estados de comprensión. Cuando pretendemos inferir los estados de comprensión en que están los sujetos, el éxito o fracaso observado debe ser tratado como es: un indicador falible del estado subyacente del sujeto.

Por todo esto, la conexión entre las observaciones y las variables del modelo de estudiante debe ser probabilística más que determinística, puesto que estas últimas formulaciones no pueden dar explicación de las inconsistencias mencionadas.

Por otra parte, de la misma tabla 5 podemos colegir que las probabilidades de respuesta correcta a algunos ítems no crecen monotónicamente con el incremento de la puntuación global del test. Estas inconsistencias no son más que el reflejo del hecho de que las tareas diseñadas no forman un conjunto completamente ordenado, por lo que la probabilidad de realizar correctamente las tareas mencionadas en función del estado latente no es una función monótona creciente. Por esto no es posible modelizar esa relación ni con la teoría clásica ni con IRT.

En consecuencia se consideró que el modelo de clase latente basado en el Lazarsfeld (1950) era el más adecuado. Cada uno de los cinco estados se consideró una clase latente. En este modelo los parámetros a estimar son las probabilidades condicionales. (Para el modelo de clase latente véase Lazarsfeld y Henry, 1968): Como resultado de la estimación se obtuvieron los datos de la tabla 6. El vector de probabilidades estimadas de cada estado de población, $P(\eta)$ viene dado por $P(\eta) = (P(\text{estado}=0), P(\text{estado}=I), P(\text{estado}=II), P(\text{estado}=III), P(\text{estado}=IV)) = (0.257, 0.227, 0.163, 0.275, 0.078)$: Con todos estos datos puede procederse al diagnóstico a partir de las respuestas observadas.

Los datos obtenidos por Siegler de las repuestas de 60 chicos comprendidos en las edades de dos a cinco años, en dos tiempos (120 respuestas) sirvieron a Mitlevy y colaboradores para inferir el estado de comprensión de los sujetos sobre el funcionamiento de la balanza .

En el ejemplo que nos ocupa, las estimaciones obtenidas del ajuste del modelo sirvieron para inicializar una red probabilística. ERGO™ (Noetic Systems Incorporated, 1992) fue el programa empleado para gestionarla. Este programa utiliza el algoritmo de actualización desarrollado por Lauritzen y Spiegelhalter que ha sido mencionado en la sección anterior. A través de un sencillo interface gráfico permite crear y gestionar redes probabilísticas muy complejas. En lo que sigue ilustraremos brevemente el funcionamiento del programa.

En la figura 7 tenemos una representación del diagrama causal realizado por ERGO™, con indicación de las probabilidades iniciales de los valores de cada

variable. Un nodo representa los estados latentes de 0 a 4, y luego por cada tipo de ítem tenemos otro nodo con dos valores, correcto e incorrecto. Las flechas indican que hay una relación de causalidad entre las variables latentes y observadas. La longitud de las barras indica la probabilidad inicial asignada a cada uno de esos valores. Los de la variable son las probabilidades a priori, $P(\eta)$, mientras que los de los ítems son las probabilidades marginales, es decir, representan la probabilidad de que si extraemos al azar un niño de esa población de esa respuesta al ítem.

¿Qué ocurre cuando obtenemos una determinada información? Por ejemplo, un alumno contesta incorrectamente a un ítem de tipo S. En la figura 8 vemos cómo se actualiza la información. En primer lugar la barra correspondiente al valor cero del ítem S se pone a 1, y su complementaria a cero. Seguidamente, se calculan los valores de $P(\eta|x)$, por aplicación del teorema de Bayes ($P(\eta|x) \propto P(x|\eta)P(\eta)$): También las probabilidades asociadas a los distintos valores de los demás ítems se actualizan. Se calculan las medias de las probabilidades de respuesta correcta esperada de los distintos estados latentes, ponderadas por las nuevas probabilidades obtenidas para cada estado.

La dinámica del sistema en una aplicación diagnóstica es muy sencilla. Cada vez que obtenemos información de un ítem, vamos precisando las probabilidades asociadas a cada estado latente. Podemos aplicar ítems hasta que decidamos que conocemos con suficiente precisión la situación real de un sujeto. Por otra parte, después de cada paso podemos decidir qué tipo de ítem aplicaremos a continuación. Para ello basta con determinar qué ítem de los restantes proporciona mayor cantidad de información.

El ejemplo presentado es naturalmente un caso en el que la estructura latente es muy sencilla. Programas como el presentado permiten manejar sistemas muy complejos con gran número de variables. Por ejemplo el mismo sistema ha sido aplicado con éxito al diagnóstico de errores en la realización de resta de fracciones, (Mislevy, 1993c, basado en un análisis cognitivo del dominio realizado por Tatsuoka, 1987, 1990; puede verse el diagrama causal en la figura 9), en el diagnóstico de destrezas de identificación de averías en sistemas hidráulicos de aviones, (Gitomer, Steinberg y Mislevy, en preparación); Murillo (1993), realiza una propuesta para operacionalizar con esta misma técnica un modelo causal en el contexto de la teoría de Sternberg sobre la comprensión verbal.

Hemos presentado en este apartado un ejemplo que representa un cierto tipo de modelización y de explotación de la información a través de un sistema experto probabilístico. Hay un gran número de trabajos en los que está presente la misma concepción diagnóstica de la evaluación. Algunos de ellos son: Marshall (1993) donde se presentan algunas sugerencias acerca de las características deseables de las nuevas alternativas evaluativas. DiBello, Stout y Roussos (1993) presentan un modelo en el que se unifican los resultados de un análisis cognitivo de cierto dominio de operaciones aritméticas y las virtualidades de los modelos psicométricos habituales. Su modelo es comentado por Samejima (1993); Biswas y Goldman (1993) desarrollan una metodología para evaluar la capacidad de resolución de

problemas en dominios complejos. Su trabajo se especifica en la evaluación de la actividad de diseño de complejos circuitos digitales de tecnología CMOS. Sobre el mismo tema trata Samejima (1992): Sheehan y Mislevy (1989) proponen un modelo para la evaluación de la alfabetización funcional basado en un análisis cognitivo de los textos en los que la información se divide en categorías jerárquicas. Dubois y Shalin (en preparación) en el que se describe la utilización de métodos cognitivos para la medida de ejecución práctica en una aplicación de evaluación de la capacidad de orientarse sobre el terreno (navegación en tierra): Un trabajo en el que se pretende realizar una integración del análisis cognitivo, los modelos psicométricos y la inteligencia artificial es el de Bennett (1993).

TABLA 6
PROBABILIDADES CONDICIONALES ESTIMADAS

Tipo de problema	Estado 0	Estado I	Estado II	Estado III	Estado IV
E	0,333*	0,973	0,883	0,981	0,943
D	0,333*	0,973	0,883	0,981	0,943
S	0,333*	0,026	0,883	0,981	0,943
CD	0,333*	0,973	0,883	0,333*	0,943
CS	0,333*	0,026	0,116	0,333*	0,943
CE	0,333*	0,026	0,116	0,333*	0,943

3. CONCLUSIONES

En resumen, podemos considerar que visto con perspectiva histórica nos encontramos con distintos escalones que nos acercan a un diagnóstico con una base psicológica más elaborada. La TCT se basa en el supuesto de que el rendimiento global en un test viene representado por una sola cantidad X. No se dice nada acerca de cómo, mediante qué mecanismos se ha llegado a generar ese valor. Sólo afirmamos que esa cantidad representa la suma de dos valores, uno de ellos la puntuación verdadera, acerca de cuya naturaleza no hay acuerdo.

La IRT representa un avance, ya que se liga un rasgo psicológico no observado al rendimiento en cada uno de los ítems. De alguna forma se intenta explicar lo que ocurre en el proceso de generación de la respuesta observada.

La nueva psicometría trata no sólo de analizar el tests en sus componentes, sino que pretende explicar la relación causal que existe entre la estructura psicológica del sujeto que responde y la respuesta que genera a los distintos ítems. Se introduce un elemento causal muy importante en la naturaleza científica de la psicometría. Puede que no sea exagerado decir que se introduce un cambio epistemológico de suma importancia.

Hasta la aparición de este nuevo enfoque teníamos dos mundos completamente

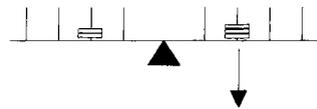
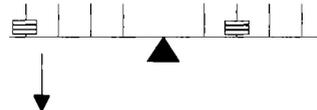
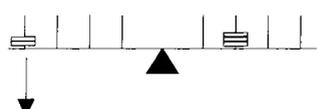
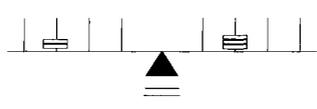
Tipo de ítem	Descripción
	Con pesos de igual magnitud v colocados a la misma distancia del fiel.
	Con pesos desiguales v colocados a la misma distancia del fiel.
	Con distancias desiguales v pesos iguales.
	Un lado tiene mayor peso v el otro mayor distancia. El lado con mayor peso descender.
	Un lado tiene mayor peso v el otro mayor distancia. El lado con mayor distancia descender.
	Un lado tiene mayor peso v el otro mayor distancia. Permanece en equilibrio.

Figura 6
Ejemplos de ítems de distintos tipos.

independientes entre sí. De un lado, el diagnóstico cognitivo era sólo posible a nivel individual. El producto de tal actividad diagnóstica es un resumen altamente informativo y complejo, con una detallada descripción del individuo estudiado, pero difícilmente comparable con el de otros sujetos. La aplicación a grandes grupos, o mediante ordenador es difícil, por no decir imposible, debido a razones de costo y de falta de formalismo. El diagnóstico en este ámbito tiene su razón de ser en el consejo individual, en la orientación individual.

Por otra parte la evaluación de rendimientos educativos en grandes grupos era sólo posible utilizando pruebas estandarizadas, cuyo resultado consiste en una sola cantidad referida a cada sujeto. El valor de esa cantidad es relativamente bajo a nivel individual, por lo poco informativo que resulta, aunque como contrapresta-

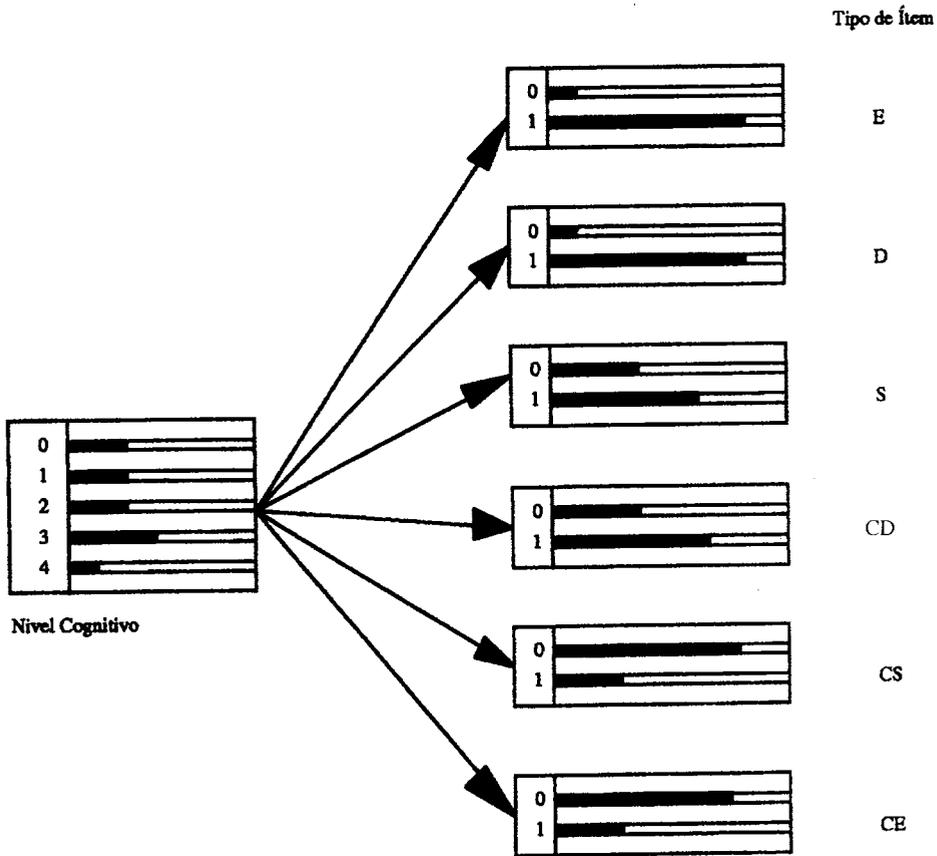


Figura 7
Estado inicial del sistema.

ción, es fácilmente comparable con los resultados de otros sujetos, y muy barato de aplicar a grandes grupos. La selección o la evaluación del sistema suele ser el objetivo final de este otro enfoque.

La nueva corriente que analizamos permite la integración de lo mejor de esas dos tendencias. Al modelo cognitivo elaborado, jerarquizado y complejo, se le dota de un grado de formalismo que permite tanto la aplicación a través de pruebas estandarizadas a grandes grupos, como mediante ordenadores y sistemas expertos, en los que las distribuciones condicionales de las respuestas asociadas a cada estado quedan reflejadas en el sistema de reglas del motor de inferencias probabilísticas.

Esta o estas nuevas teorías de los tests, admite, o incluso impulsa y necesita, modelos más complejos y elaborados de los sujetos que aprenden. Esto tiene una doble consecuencia. Por una parte la psicometría deja de ser una sofisticada elabo-

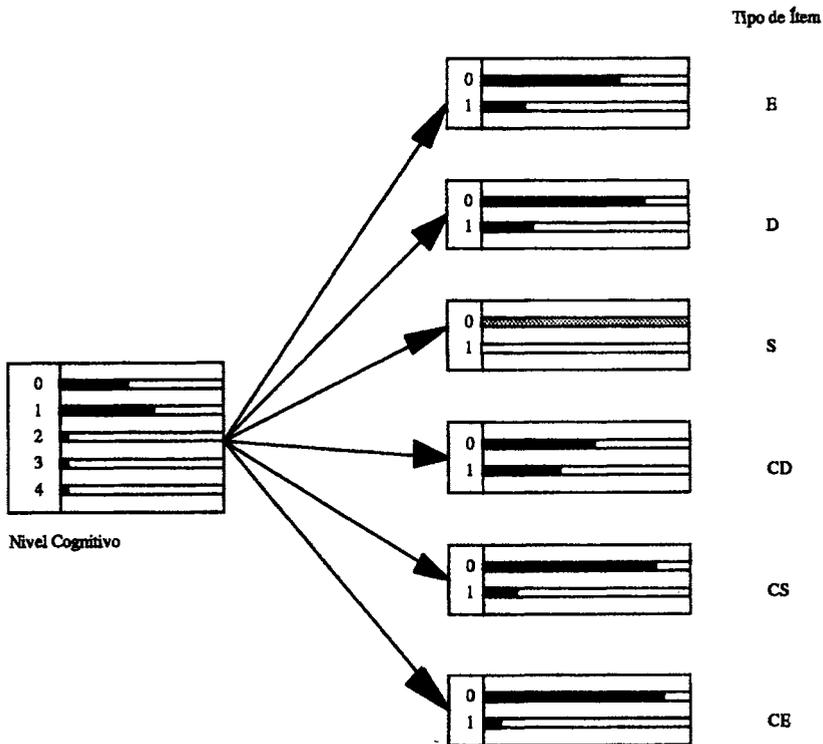


Figura 8
Estado del sistema tras una respuesta.

ración formal, un alto edificio con una gran estructura matemática y una estrechísima base psicológica. La base psicológica comienza a crecer, a la par que se integra en la estructura matemática. Los modelos formulados en la TCT y la IRT son, en términos de BUNGE, modelos **fenomenológicos**, es decir, modelos en los que lo que se pretende es la simple reproducción de lo manifiesto del fenómeno, del epifenómeno. La predicción a corto plazo o en condiciones muy limitadas es lo máximo que permiten este tipo de modelos. Son modelos de caja negra, en los que no aparece ningún elemento que pueda definirse claramente como elemento causal.

En cambio, los modelos que en la nueva teoría de los tests se propugnan, no hacen una simple descripción de lo que aparece en el proceso de generación de la respuesta observable; la puntuación en el test, en la TCT, y la puntuación en el ítem, en la IRT. Ahora sin embargo se pretende modelizar al propio individuo que genera la respuesta en cuanto a los mecanismos que pone en juego para producirla. Se pretende elaborar una representación del «mecanismo» que actúa cuando la conducta observada se genera. Se suponen unos estados por los que el sujeto modelizado

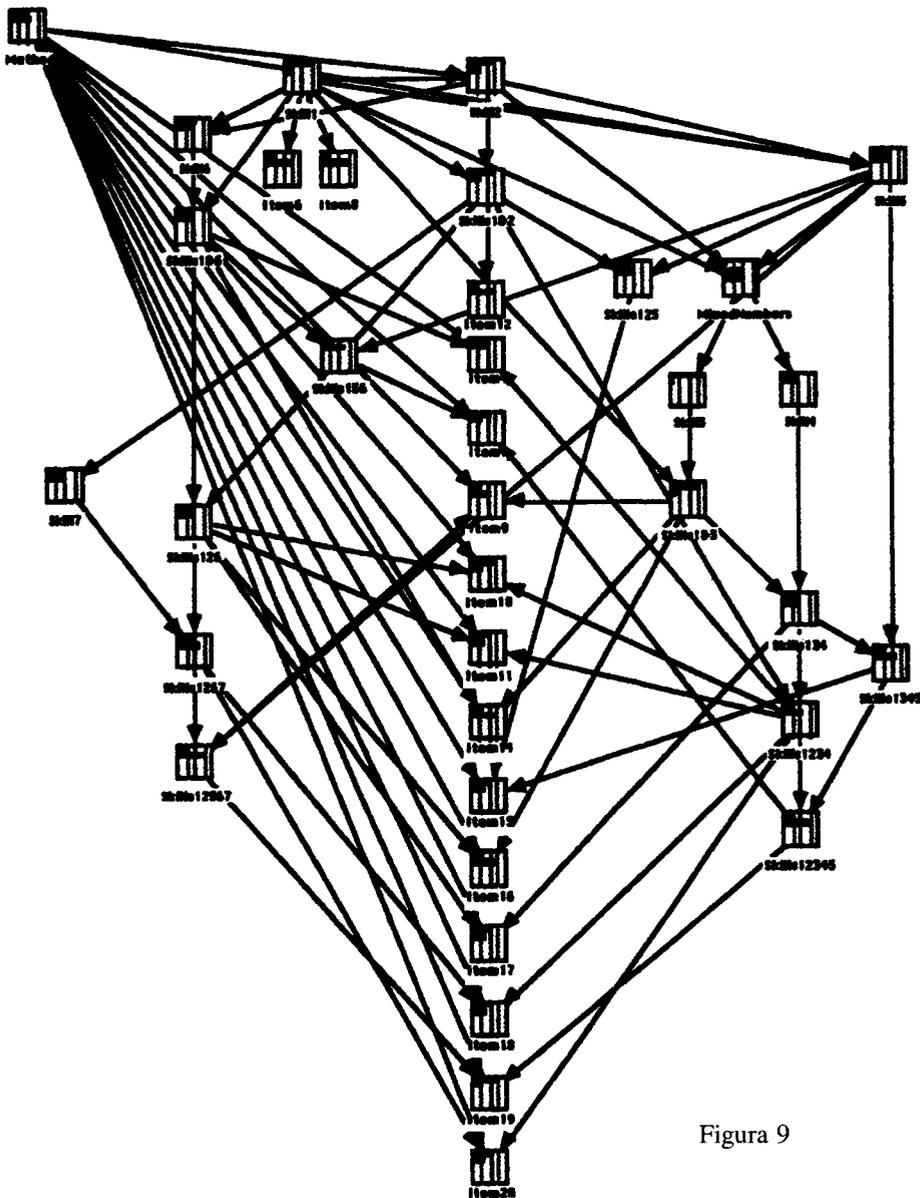


Figura 9

puede pasar, ligando a cada uno de esos estados una distribución de probabilidad de las distintas respuestas posibles. Se introduce un elemento de causalidad, o al menos la posibilidad de que estos modelos lo incluyan. El alumno que responde ya no es sólo un valor en una variable continua, sino que su complejidad psicológica puede ser modelizada, y nuestros conocimientos sobre ella incluidos en el modelo psico-

métrico. ¿Quiere eso decir que los «esquemas», son más «reales» que los rasgos? La respuesta es que no. Los esquemas también son construcciones auxiliares que permiten inducir independencia local o condicional, y por lo tanto construir esquemas de explicación causal. Sin embargo esa construcción auxiliar es más sofisticada que sus predecesoras, y por eso es que la intención asociada a su formulación es precisamente la de ser isomórfica con la realidad. Por esto, y también usando la terminología de BUNGE, se trata de modelos, al menos localmente, **representacionales**, es decir, explicativos y causales. Desde este punto de vista la superioridad epistemológica de esta corriente es evidente. Pero además lo anterior implica que **el modelo que guía el aprendizaje y el que guía la evaluación son la misma cosa**. Esta capacidad de integrar en un sólo conjunto el aprendizaje y su evaluación supone un avance no sólo epistemológico sino de orden práctico.

Con respecto al nivel de generalidad de los modelos cognitivos y sus repercusiones epistemológicas, es necesario hacer algunas precisiones. En psicología cognitiva los modelos son en su mayoría modelos locales. Tal vez sea esa característica la que ha producido el malentendido de la atribución de un carácter cualitativo a los mismos. Si aceptamos que el rasgo más sobresaliente del paradigma cuantitativo es la pretensión de desarrollar explicaciones causales, está claro que los modelos a los que nos hemos referido lo son. Pero posiblemente queda mejor aclarada la idea si pensamos que las explicaciones pueden tener dos niveles distintos de generalidad: respecto de los sujetos y respecto de los contenidos. El tipo de discurso desarrollado dentro del paradigma cualitativo se asienta sobre el supuesto de que el dato fundamental no es una realidad objetiva externa al sujeto que conoce, sino más bien la interpretación que dicho sujeto hace del contexto en el que se encuentra. En el paradigma cognitivo, es cierto que interesa la arquitectura de las representaciones que construye cada sujeto, pero se parte del supuesto de la existencia de unos mecanismos generales de la cognición que son comunes a todos los sujetos. Por tanto se da generalidad de la validez de las teorías en lo que se refiere a los sujetos. La localidad a que nos referimos se da en cuanto a los contenidos de la cognición. Los modelos son específicos de los dominios de contenido, no de los sujetos. Desde este punto de vista por tanto podemos atribuir a estos modelos capacidad de explicación causal, si bien llamamos la atención sobre el hecho de que no se deducen de un modelo general de la cognición, sino que se construyen específicamente para explicar fenómenos locales.

La localidad de estos modelos plantea algunas cuestiones importantes. Es innegable que en la resolución de problemas diagnósticos concretos las teorías locales y sus modelos correspondientes pueden resultar suficientemente útiles como para justificar su existencia. El diagnóstico es una actividad práctica, y como tal, no necesita de otros niveles de generalidad. Sin embargo desde un punto de vista práctico también es cierto que entonces es necesario multiplicar los esfuerzos de resolución de problemas diagnósticos.

Por otra parte desde el punto de vista de la construcción de una ciencia es necesaria la existencia de un marco general de explicación. En una situación ideal,

una teoría general de la cognición podría producir, al enfrentarse con las situaciones planteadas en el contacto con los dominios determinados de conocimientos, modelos específicos que serían casos particulares del primero. Hoy no tenemos todavía esa jerarquía. Pero no hay que descartar que la profusión de teorías locales sirva para allanar el camino hacia la construcción de un sistema general de referencia de alto nivel de abstracción.

El futuro de este paradigma está ligado, probablemente, a su capacidad para generar esa teoría general de la cognición a que nos hemos referido y derivada o complementaria de ella, una teoría general del aprendizaje, no ligadas por tanto, específica y exclusivamente a contextos limitados.

Por otra parte se consigue una integración para muchos insospechada. Ciertos modelos y teorías cognitivas acerca de la inteligencia y el aprendizaje, cuya naturaleza y estructura les acercaba más al mundo de lo cualitativo y por ende los hacía declarados enemigos de la utilización de pruebas estandarizadas para el diagnóstico, pueden ahora, sin embargo, proporcionar la base psicológica para un nuevo tipo de diagnóstico. Y esto en definitiva no hace más que subrayar el malentendido mencionado que colocaba a la ciencia cognitiva del lado de lo cualitativo, o mejor, enfrente de lo cuantitativo. Pero la distinción fundamental entre uno y otro paradigma radica en que el segundo se trata de desarrollar explicaciones causales. Y desde ese punto de vista, gran parte de los modelos y teorías formuladas en el ámbito de las ciencias cognitivas son, indudablemente, explicativas, y por tanto cuantitativas, a pesar de que sus sistemas de representación se basen en categorías a veces no ordenadas, o parcialmente ordenadas, y por tanto no lleguen a utilizar sistemas de numeración.

La que hemos presentado es una línea de investigación que está comenzando a florecer en estos momentos. No es posible saber si en el futuro se convertirá en la corriente principal en el área de la psicometría y el diagnóstico. No sabemos qué otros desarrollos pueden más adelante obligarnos a mirar interesados en otras direcciones. Lo que es indudable es que estos postulados obligan a plantearse importantes cuestiones acerca de la relación entre la psicometría, y el aprendizaje y su diagnóstico. Tal vez caigamos en la cuenta de que ya no es posible hablar del diagnóstico educativo sin referirnos a la psicometría, ni a ésta última sin mencionar al primero, ya que ambas áreas son básicamente procesos de inferencia en orden a la mejora y la intervención educativa.

4. REFERENCIAS BIBLIOGRÁFICAS

- BEJAR, I. I. (1993): A generative approach to psychological and educational measurement. En N. FREDERIKSEN, R. J. MISLEVY e I. I. BEJAR (eds.): *Test Theory for a new Generation of Tests*. Lawrence Erlbaum, Hillsdale, N. J., 323-357.
- BEJAR, I. I. (1988): «A sentence-based automated approach to the assessment of writing: A feasibility study» *Machine-Mediated Learning*, 2, 321-332.
- BEJAR, I. I., CHAFFIN, R. y EMBRETSON, S. (1991): *Cognitive and psychometric analysis of analogical problem solving*, Springer-Verlag, N.Y.

- BENNETT, R. E. (1993): Toward intelligent assessment: an integration of constructed-response testing, artificial intelligence and model-based measurement. En N. FREDERIKSEN, R. J. MISLEVY e I. I. BEJAR (eds.): *Test Theory for a new Generation of Tests*, Lawrence Erlbaum, Hillsdale, N. J., 99-123.
- BENNETT, R. E., WARD, W. C., ROCK, D. A. y LAHART, C. (1990): Toward a framework for constructed-response ítems. Documento nº RR-90-7. Educational Testing Service. Princeton, N.J.
- BERGAN, J. R. y STONE, C. A. (1985): «Latent class models for knowledge domains» *Psychological Bulletin*, 98, 166-184.
- BIREMBAUM, M. & TATSUOKA, K. K. (1987): Open-ended versus multiple-choice response formats - It does make a difference for diagnostic purposes. *Applied Psychological Measurement*, 11, 385-395.
- BISWAS, G. y GOLDMAN, S. (1993): Assessing design activity in complex CMOS circuit design. Documento presentado en Conference on Alternative Diagnostic Assessment, The university of Iowa, 6-7 mayo.
- BOCK, R. D. (1972): «Estimating item parameters and latent ability when responses are scored in two or more nominal categories». *Psychometrika*, 37, 29-51.
- BRAUN, H., CARLSON, S. y BEJAR, I. I. (1989): *Psychometric foundations of testing based on patient management problems* (RM-89-2) Educational Testing Service, Princeton.
- BROWN, J. S. y BURTON, R. R. (1978): «Diagnostic models for procedural bugs in basic mathematical skills» *Cognitive Science*, 2, 155-192.
- BUTTERFIELD, E. C., NIELSEN, D., TANGEN, K. L. y RICHARDSON, M. B. (1985): Theoretically-based psychometric measures of inductive reasoning. En EMBRESTON, S. (ed.): *Test design: developments in psychology and psychometrics*. Academic Press, Orlando, pp. 77-148.
- CAMPBELL, D. T. y FISKE, D. W. (1959): «Convergent and discriminant validation by the multitrait-multimethod matrix» *Psychological Bulletin*, 56, 81-105.
- CARRETERO, M., POPE, M., SIMONS, R. J. y POZO, I. (eds.) (1991): *Learning and Instruction*. European Research in an International Context, Vol. III, Pergamon, Oxford.
- CHAFFIN, R. y HERRMANN, D. J. (1987): Relation element theory: A new account of the representation and processing of semantic relations. En GORFEIN, D. y HOFFMAN, R. (eds.): *Memory and Learning: The Ebbinghaus centennial conference*, Erlbaum, Hillsdale, N.J.
- CHOMSKY, N. (1965): *Aspects of the theory of syntax*, MIT Press, Cambridge MA.
- CORBALLIS, M. C. (1982): Mental rotation: Analysis of a paradigm. En POTE GAL, M. (ed.): *Spatial abilities: developmental and psychological foundations*, 173-198, Academic Press, N.Y.
- DIBELLO, L., STOUT, W. y ROUSSOS, L. (1993): Unified cognitive/psychometric diagnosis foundations and applications. Documento presentado en Conference on Alternative Diagnostic Assessment, The university of Iowa, 6-7 mayo.
- DUBOIS, D. y SHALIN, V. (1993): Adapting cognitive methods to real world objectives: An application to job Knowledge testing. Documento presentado en «Conference on Alternative Diagnostic Assessment», The university of Iowa, Mayo.
- DUNCAN, J. y HUMPHREYS, G. W. (1989): «Visual search and stimulus similarity» *Psychological Review*, 96, 433-458.
- ELSHOUT (1985): Problem Solving and Education. Documento presentado en la reunión de la «European Association for Research on Learning and Instruction». Bélgica.
- EMBRESTON, S. (1985): Multicomponent latent trait models for test design. En EMBRESTON, S. (ed.): *Test design: developments in psychology and psychometrics*. Academic Press, Orlando.
- EMBRESTON, S. (1993): Psychometric models for learning and cognitive processes. En N. FREDERIKSEN, R. J. MISLEVY e I. I. BEJAR (eds.): *Test Theory for a new Generation of Tests*. Lawrence Erlbaum, Hillsdale, N.J., 125-150.

- FREDERIKSEN, N., MISLEVY, R. J. y BEJAR, I. I. (eds.): (1993): *Test Theory for a new Generation of Tests*, Lawrence Erlbaum, Hillsdale, N.J.
- FREDERIKSEN, N., GLASER, R., LESGOLD, A. y SHAFTO, M. G. (eds.) (1990): *Diagnostic monitoring of skill and Knowledge acquisition*, Lawrence Erlbaum, Hillsdale, N.J.
- FULTON, S. L. y PEPE, C. O. (1990): An introduction to model-based reasoning. *AI Expert*, 48-55.
- GAVIRIA SOTO, J. L. (En preparación): Un modelo del tiempo de respuesta para predecir la habilidad latente. Universidad Complutense.
- GELMAN, R. y GALLISTEL, R. (1978): *The child's understanding of number*. Harvard University Press, Cambridge MA.
- GITOMER, D. H., STEINBERG, L. S. y MISLEVY, R. J. (en preparación): Diagnostic assessment of troubleshooting skill in an intelligent tutoring system.
- GITOMER, D. H. y YAMAMOTO, K. (1991): Performance Modeling that integrates latent trait and class theory. Research Report RR-91-1, Educational Testing Service, Princeton, N.J.
- GLASER, R., LESGOLD, A. y LAJOIE, S. (1987): Toward a cognitive theory for the measurement of achievement. En RONNING, R., GLOVER, J., CONOLEY, J. C. y WITT, J. (eds.): *The influence of cognitive psychology on testing and measurement: The Buros-Nebraska Symposium on measurement and testing*, Vol. 3, Erlbaum, Hillsdale, N.J.
- GOLDIN, G. A. y McCLINTOCK, C. E. (eds.) (1984): *Task variables in mathematical problem solving*. Franklin Institute Press, Philadelphia.
- HIVELY, W., PATERSON, H. L. y PAGE, S. H. (1968): «A universe-defined system of arithmetic tests». *Journal of Educational Measurement*, 5, 275-290.
- HAERTEL, E. H. y WILEY, D. E. (1993): Representations of ability structures: implications for testing. En N. FREDERIKSEN, R. J. MISLEVY e I. I. BEJAR (eds.): *Test Theory for a new Generation of Tests*, Lawrence Erlbaum, Hillsdale, N.J., 359-384.
- HAERTEL, E. H. (1990): «Continuous and discrete latent structure models for item response data» *Psychometrika*, 55, 3, 477-494.
- JARJOURA, D. y BRENNAN, R. L. (1982): «A variance components model for measurement procedures associated with a table of specifications». *Applied Psychological Measurement*, 6, 161-171.
- JOHNSON-LAIRD, P. N., BYRNE, R. M. J. y TABOSSI, P. (1989): «Reasoning by model: The case of multiple quantification». *Psychological Review*, 96, 658-673.
- KATZ, B. (1988): *Using English for indexing and retrieving* (A.I. Memo n° 1.096) Massachusetts Institute of Technology, Cambridge, MA.
- KIERAS, D. E. (1990): The role of cognitive simulation models in the development of advanced training and testing systems. En FREDERIKSEN, N., GLASER, R., LESGOLD, A. y SHAFTO, M. G. (eds.) (1990): *Diagnostic monitoring of skill and Knowledge acquisition*, (51-73) Lawrence Erlbaum, Hillsdale, N.J.
- KLASSEN, D. y O'CONNOR, W. A. (1987): Predicting violence in mental patients: Cross-validation of an actuarial scale. Documento presentado en el Annual Meeting of the American Public Health Association, New Orleans.
- KOCH, W. R. (1983): «Likert scaling using the graded response latent trait model» *Applied Psychological Measurement*, 7, 15-32.
- KOLEN, M. J. y HARRIS, D. J. (1987): A multivariate test theory model based on item response theory and generalizability theory. Presentado en Annual Meeting de AERA.
- LAZARSFELD, P. F. y HENRY, N. W. (1968): *Latent Structure Analysis* Houghton-Mifflin, Boston.
- LAURITZEN, S. L. & SPIEGELHALTER, D. J. (1988): Local computations with probabilities on graphical structures and their application to expert systems. (Le sigue un debate): *Journal of the Royal Statistical Society, Series B*, 50, 157-224.

- LESGOLD, A., IVILL-FRIEL, J. y BONAR, J. (1989): Toward intelligent systems for testing. En RESNICK, L. R. (ed.): *Knowing, learning, and instruction: Essays in honor of Robert Glaser* (337-360) Lawrence Erlbaum, Hillsdale N.J.
- LINN, R. L. (1989) (ed.): *Educational Measurement*, (3ª edic.), Macmillan, N.Y.
- LOHMAN, D. F. e IPPEL, J. (1993): Cognitive diagnosis: From Statistically based assessment toward theory-based assessment. En N. FREDERIKSEN, R. J. MISLEVY e I. I. BEJAR (eds.): *Test Theory for a new Generation of Tests*, 41-71. Lawrence Erlbaum, Hillsdale, N.J.
- LORD, F. M. y NOVICK, M. R. (1968): *Statistical Theories of Mental Test Scores*. Reading, Mass.: Addison-Wesley.
- MARSHALL, S. P. (1990): Generating good ítems for diagnostic tests. En FREDERIKSEN, N., GLASER, R., LESGOLD, A. y SHAFTO, M. G. (eds.) (1990): *Diagnostic monitoring of skill and Knowledge acquisition*, Lawrence Erlbaum, Hillsdale, N.J.
- MARSHALL, S. P. (1993): Some suggestions for alternative assessment. Documento presentado en «Conference on Alternative Diagnostic Assessment», The university of Iowa, 6-7 mayo.
- MARSHALL, S. P. (1993): Assessing schema knowledge. En N. FREDERIKSEN, R. J. MISLEVY e I. I. BEJAR (eds.): *Test Theory for a new Generation of Tests*, Lawrence Erlbaum, Hillsdale, N.J., 155-180.
- MASTERS, G. N. (1982): «A Raxh model for partial credit scoring» *Psychometrika*, 47, 149-174.
- MASTERS, G. N. y WRIGHT, B. D. (1984): «The essential process in a family of measurement models» *Psychometrika*, 49, 269-272.
- MCDONALD, R. P. (1981): The dimensionality of tests and ítems. *British Journal of Mathematical and Statistical Psychology*, 34, 100-117.
- MCDONALD, R. P. (1985): *Factor Analysis and Related Methods*. Lawrence Erlbaum, Hillsdale, N.J.
- MESSICK, S. (1989): Validity. En LINN, R. L. (1989) (ed.): *Educational Measurement*, (3ª edic.), Macmillan, N.Y., 13-103.
- MESSICK, S. (1992): The interplay of evidence and consequences in the validation of performance Assessment. Research Report RR-92-39, Educational Testing Service, Princeton, N.J.
- MILLER, P. L. (1984): *A critiquing approach to expert computer advice: Attending*. Kaufmann, Palo Alto, CA.
- MISLEVY, R. J. (1991): A framework for studying differences between multiple-choice and free-response test ítems. Educational Testing Service. Princeton, N.J.
- MISLEVY, R. J. (1993a): Foundations of a new test theory. En N. FREDERIKSEN, R. J. MISLEVY e I. I. BEJAR (eds.): *Test Theory for a new Generation of Tests*. Lawrence Erlbaum, Hillsdale, N.J., 19-39.
- MISLEVY, R. J. (1993b): Probability-Based inference in Cognitive Diagnosis. Documento presentado en la Conference on alternative diagnostic assessment. Universidad de Iowa, 6-7 de mayo.
- MISLEVY, R. J. (1993c): Tests Theory Reconcepted. Documento presentado en la reunión anual de AERA.
- MISLEVY, R. J., YAMAMOTO, K. y ANACKER, S. (1990): Toward a test theory for assessing student understanding. Educational Testing Service, Princeton, N.J.
- MURAKI, E. (1989): *Manual for Parscale Program*. Scientific Software. Fairplay, Co.
- MURILLO TORRECILLA, F. J. (1993): Aplicación de los nuevos enfoques de teoría de la medida en la construcción de una prueba diagnóstica de comprensión verbal. Comunicación presentada en el VI Seminario de Modelos de Investigación Educativa, Madrid, 23-25 de septiembre.
- NEWELL, A. y SIMON, H. (1972): *Human Problem-solving* Englewood Cliffs, Prentice-Hall.
- ORASANU, J. y FISCHER, U. (1992): Team cognition in the cockpit: linguistic control of shared problem-solving. En *Proceedings of the 14th Annual Meeting of the Cognitive Science Society*. Hillsdale, N.J.

- PEARL, J. (1988): Probabilistic reasoning in intelligent system: Networks of plausible inference. San Mateo, CA: Kaufmann.
- RECKASE, M. D. (1979): *Unifactor Latent Trait Models applied to multifactor tests: Results and implications*. Journal of Educational Statistics, 4, 207-230.
- RESNICK, L. B. y RESNICK, D. P. (1989): Assessing the thinking curriculum: new tools for educational reform. En GIFFORD, B. R. y O'CONNOR, M. C. (eds.): *Future Assessment: changing assessments: alternative views of aptitude, achievement, and instruction*. Kluwer Academic Publishers, Boston, 37-75.
- ROID, G. y HALADYNA, T. (1982): *A technology for test-item writing*, Academic Press, N.Y.
- RUIZ PRIMO, A.; BAXTER, G. y SHALVESON, R. (1993): «On stability of performance assessment». Journal of Educational Measurements. 30, 1, 41-53.
- SAMEJIMA, R. (1969): «Estimation of latent ability using a response pattern of graded scores». Psychometric Monograph, nº 17. Psychometric Society.
- SAMEJIMA, F. (1993): A design of cognitive diagnosis. Documento presentado en «Conference on Alternative Diagnostic Assessment», The university of Iowa, 6-7 mayo.
- SAMIJIMA, R. (1993): A cognitive diagnosis method using latent trait models. Documento presentado en «Conference on Alternative Diagnostic Assessment», The university of Iowa, 6-7 mayo.
- SHAFER, G., & SHENOY, P. (1988): Bayesian and belief-function propagation. Working paper 121. Lawrence, KS: Schools of Business, University of Kansas.
- SHEEHAN, K. y MISLEVY, R. (1989): Integrating cognitive and psychometric models to measure document literacy. Documento nº (RR-89-51-ONR) Educational Testing Service. Princeton, N.J.
- SHEPARD, L. (1991): «Psychometricians beliefs about learning» *Educational Researcher*, 20, 7, 2-16.
- SIEGLER, R. S. (1981): Developmental sequences within and between concepts. Monograph of the Society for Research in Child Development, 46.
- SNOW, R. E. y LOHMAN, D. F. (1989): Implications of cognitive psychology for educational measurement. En LINN, R. L. (ed.): *Educational Measurement*, (3ª edic.), Macmillan, N.Y., 263-331.
- SNOW, R. E. y LOHMAN, D. F. (1993): Cognitive psychology, new test design and new test theory: An introduction. En N. FREDERIKSEN, R. J. MISLEVY e I. I. BEJAR (eds.): *Test Theory for a new Generation of Tests* Lawrence Erlbaum, Hillsdale, N.J., 1-17.
- STERNBERG, R. J. (1984): What cognitive psychology can and cannot do for test development. En B. S. PLAQUE (Ed.): *Social and technical issues in testing: Implications for test construction and usage*, Lawrence Erlbaum, Hillsdale, N.J., 39-60.
- TATSUOKA, K. K. (1990): Toward an integration of Item-Response Theory and cognitive error diagnosis. En FREDE RIKSEN, N., GLASER, R., LESGOLD, A. y SHAFTO, M. G. (eds.) (1990): *Diagnostic monitoring of skill and Knowledge acquisition*, Lawrence Erlbaum, Hillsdale, N.J. 453-488.
- THISSEN, D. (1993): Repealing rules that no longer apply to psychological measurement. En N. FREDERIKSEN, R. J. MISLEVY e I. I. BEJAR (eds.): *Test Theory for a new Generation of Tests*, Lawrence Erlbaum, Hillsdale, N.J., 79-97-
- THISSEN, D. M. y STEINBERG, L. (1986): «A taxonomy of item response models». *Psychometrika*, 51, 567-577.
- TRYON, R. C. (1957): «Reliability and behavior domain validity: reformulation and historical critique» *Psychological Bulletin*, 54, 229-249.
- WANG, M. (1985): *Fitting a Unidimensional Model to Multidimensional Item Response Data: The Effects of Latent Space Misspecification on the Application of IRT*. Research Proposal, The University of Iowa.
- WANG, M. (1987): *Estimation of ability parameters from response data to ítems that are precalibrated with a unidimensional model*. Comunicación presentada en la reunión anual de la American Educational Research Association, Wasington, D.C.
- WHITELY, S. E. (1980): *Multicomponent latent trait models for ability tests*. *Psychometrika*, 45, 479-494.

YAMAMOTO, K. y GITOMER, D. H. (1993): Application of a HYBRID model to a test of cognitive skill representation. En N. FREDERIKSEN, R. J. MISLEVY e I. I. BEJAR (eds.): *Test Theory for a new Generation of Tests*, Lawrence Erlbaum, Hillsdale, N.J. 275-295.

PONENCIA IV

ANÁLISIS DE DATOS CUALITATIVOS EN LA INVESTIGACIÓN SOBRE LA DIFERENCIACIÓN EDUCATIVA

por

Eduardo García Jiménez, Javier Gil Flores, Gregorio Rodríguez Gómez
Universidad de Sevilla

EL DATO CUALITATIVO EN EL ESTUDIO DE LA DIFERENCIACIÓN EDUCATIVA

El estudio del hecho diferencial en educación, y por consiguiente la formulación de propuestas educativas que respeten las diferencias, se enfrenta en la actualidad con problemas que no pueden abordarse desde el mero análisis de las diferencias individuales. La comprensión de cuestiones como el rendimiento académico del estudiante no puede hacerse sin incluir —junto a los elementos tradicionalmente considerados— al centro escolar como unidad de análisis y, por ende, a aspectos tales como el clima y la cultura escolar o el intercambio con el entorno y el medio social. Del mismo modo, la intervención diferenciada requiere hoy propuestas que consideren la acción educativa en su conjunto, incluyendo planteamientos organizativos y pedagógicos: programas formales con estrategias de enseñanza adaptadas; actividades educativas no formales e informales; y esquemas organizativos que mejoren el clima de relaciones en el aula y el centro, y faciliten la implicación de la comunidad.

La explicación de estos nuevos y viejos problemas o la mejora de esas situaciones educativas, posiblemente esté reclamando un nuevo marco conceptual y metodológico desde el que entender la problemática de la diferenciación educativa. Así, algunos investigadores están proponiendo aproximaciones conceptuales que se fundamentan en planteamientos sociológicos o antropológicos y aproximaciones metodológicas que se apoyan en planteamientos de investigación que podríamos identificar como cualitativos. Bartolomé Pina (1992), al analizar el problema de las diferencias culturales, reclama un tipo de investigación cualitativa que, sobre todo,

asegure que las categorías culturales a introducir en la investigación sean comprendidas por los sujetos, que tenga en cuenta el pluralismo cultural y lingüístico, que triángule datos, observaciones y métodos en diferentes contextos; insistiendo en la adecuación de la observación persistente y prolongada como medio de acercamiento a una comprensión más profunda del fenómeno multicultural. En el campo de la integración de los sujetos con necesidades educativas especiales, Bodgan (1986) reclama una aproximación sociológica y el uso de la investigación cualitativa como medios para abordar el análisis del contexto donde se desarrollan los procesos educativos; y Hegarty (1988) considera que procesos como la organización de los servicios de educación especial o el desarrollo de los proyectos de integración son excesivamente complejos como para ser controlados, existiendo variables claves que no pueden ser cuantificadas y dificultades para formular hipótesis en los términos adecuados para la utilización de diseños experimentales.

La idoneidad de las estrategias de investigación cualitativa para dar respuesta a algunos de los problemas que tiene hoy planteados la diferenciación educativa, está también refrendada por estudios empíricos realizados desde campos tan diversos como el bilingüismo, las diferencias debidas al sexo, la educación multicultural, la atención de los déficits socioculturales o el estudio de las diferencias individuales en razón de los procesos cognitivos.

Desde la perspectiva de la educación multicultural, en una revisión realizada sobre campos como la integración del contenido, los procesos de construcción del conocimiento, la reducción del prejuicio, la pedagogía de la equidad y el desarrollo de la cultura escolar, se recogen diferentes estudios que utilizan enfoques etnográficos al aplicar técnicas tales como entrevistas, documentos, observaciones...etc. (Banks, 1993):

Dentro de la educación infantil y en el estudio de problemas relativos a la intervención temprana, la educación de los padres, la continuidad y coherencia entre ciclos y niveles, la evaluación del contexto en que tiene lugar la interacción propia de todo proceso de enseñanza-aprendizaje, las áreas de contenido y las políticas educativas en general, se están planteando formas de indagación intrapersonales y centradas en el contexto, tales como estudios de caso, observaciones estructuradas, entrevistas... etc. (Bowman, 1993): En el campo de la educación bilingüe, autores como Fillmore y Valadez (1986) consideran que los métodos de investigación etnográficos resultan adecuados para estudiar los modelos de interacción social y lingüística entre los alumnos y los profesores en clases bilingües. Asimismo, en los estudios sobre compensación educativa se destaca la importancia de la investigación-acción como un componente necesario en la planificación, ejecución y evaluación de los programas de educación compensatoria (Espín, 1991).

Las investigaciones sobre las diferencias individuales en función del sexo, realizadas por autores como Alberdi (1985, 1986), Alberdi y Escario (1987), utilizan técnicas cualitativas como los grupos de discusión para analizar las actitudes y comportamientos de los profesores; o autores como Brullet y Subirats (1990) abogan por la necesidad de contar con la «investigación-acción» como medio funda-

mental en el proceso de transformación de la realidad educativa sexista; a lo que habría que añadir la necesidad de contar con medios y estrategias que nos acerquen a un mayor conocimiento de lo que sucede en las aulas (Caselles, 1991).

La investigación sobre la eficacia escolar o sobre la mejora de las escuelas, ya desde su inicio, ha utilizado estrategias metodológicas de corte cualitativo, recordemos en este sentido que uno de los primeros trabajos realizados fue el estudio de casos llevado a cabo por parte de Weber (1971): En los últimos años se viene reclamando la necesidad de utilizar estrategias metodológicas que saquen a la luz la especificidad de las características idiosincráticas de cada centro educativo, para adaptar su educación a las características contextuales donde se inserta (Rodríguez, 1992).

Esta progresiva inclusión de dimensiones cualitativas en la investigación sobre la diferenciación educativa, está propiciando nuevas formas de evidencia (datos) más contextualizadas y holísticas, próximas a la realidad y a la práctica, y que ofrecen una información más rica, específica y matizada. Este tipo de evidencias que, en adelante, identificaremos como datos cualitativos constituyen la base sobre la que se sitúa el desarrollo de esta ponencia.

El dato cualitativo puede definirse como una elaboración primaria que nos informa acerca de la existencia de una realidad, sus propiedades o el grado en que éstas se manifiestan. Dada su naturaleza de aproximación inicial a la realidad, el dato cualitativo ofrece una información básica con un relativamente bajo nivel de estructuración teórica. De hecho, los datos cualitativos se nos presentan como simples registros o asientos que dan fe de una realidad o informan de sus cualidades¹.

Mientras el número indica la cuantía o grado de las diferencias educativas, el dato cualitativo permite explicar la diferenciación desde un contexto holístico, que aborda la realidad en su conjunto sin separar los elementos que la explican. El dato cualitativo supone partir de una realidad educativa inicial sin apenas modificaciones; es, por tanto, una elaboración de bajo nivel, pues al investigador puede no interesarle modificar la realidad hasta no conocerla tal como era inicialmente. El dato cualitativo, además, realiza una explicación contextual de las diferencias permitiendo profundizar en los matices, contradicciones, anécdotas, peculiaridades, particularidades de personas, grupos, instituciones, etc.

El dato cuantitativo establece una diferenciación parcial, en función de un número limitado de características medidas en un momento preciso de la vida del sujeto; el dato cualitativo podría permitir una diferenciación global, en función de un elevado número de rasgos conocidos del sujeto, el grupo, la institución, su historia y trayectoria a lo largo del tiempo, significado y consecuencias de las características o de su contexto. El dato cuantitativo conduce a propuestas de adaptación educativa

¹ El dato cuantitativo por contra es altamente deductivo, una vez codificado no hay manera de recuperar la manifestación conductual original. Es como si el dato hubiese perdido la memoria de la conducta, situación y contexto de los que procede.

basadas en determinados tipos y rasgos previamente identificados, sobre la base de un individuo o grupo normativos (adaptación de carácter nomotético); el dato cualitativo lleva a propuestas de adaptación educativa acordes a las necesidades de cada sujeto concreto, teniendo en cuenta todas sus características individuales y su contexto socioeducativo (adaptación de carácter ideográfico).

Precisemos la definición de dato cualitativo a partir de dos de sus componentes, el nivel de elaboración y la forma de registro, y analicemos cómo se interpreta la naturaleza de este tipo de datos desde diferentes posiciones de investigación.

Desde determinados programas de investigación o «colegios invisibles», que cabría identificar con planteamientos positivistas o neopositivistas, el bajo nivel de elaboración teórica presente en los datos cualitativos responde a la ausencia misma de un cuerpo teórico suficientemente contrastado, de suerte que el dato cualitativo es entendido como una manifestación del nivel protocientífico en que se encuentra determinado conocimiento de la realidad. Desde esta perspectiva, si existen datos cualitativos es porque nuestra capacidad explicativa no nos permite una formulación de problemas más fundamentales, un acercamiento más selectivo a la realidad o mediciones más precisas de un objeto de conocimiento.

No conocemos los datos físicos más directa u objetivamente que los datos sociales, excepto en la medida en que hemos desarrollado instrumentos y símbolos más adecuados para registrar los primeros (Lundberg, 1968. Citado en Tejedor, 1988b:15):

Desde otros programas de investigación, que podrían identificarse con esquemas de trabajo antropológicos o sociológicos, el bajo nivel de elaboración teórica presente en los datos cualitativos es una consecuencia de la posición de partida adoptada por el propio investigador: la construcción del conocimiento es una tarea social o una consecuencia de la interacción entre un sujeto y la unidad social de la que forma parte, no el resultado de una elaboración científica previa. El conocimiento que se genera desde este tipo de programas se presenta a menudo bajo la forma de esquemas explicativos o interpretativos, generalmente de carácter operativo o práctico, que pueden llegar a tomar la forma de teorías de la acción o teorías operativas, es decir, aquellas cuya intencionalidad es la de servir de guías para el diseño de la acción educativa.

El registro de los datos cualitativos permite que la información contenida en los mismos pueda ser efectivamente manipulada y comunicada para responder a un problema de investigación. Para ello pueden emplearse distintos soportes físicos o formas de lenguaje que permitan registrar y expresar la información elaborada. Los datos cualitativos pueden quedar registrados en forma textual: transcripciones de entrevistas, notas de campo, diarios, y otros documentos. Incluso cuando los datos aparecen elaborados inicialmente en forma de imágenes o sonidos, suelen ser traducidos a una forma textual para proceder a su análisis. No obstante, es habitual que los datos cualitativos recogidos inicialmente en forma de textos puedan ser expresa-

dos con posterioridad mediante otra forma de lenguaje. El recuento de la aparición de determinados términos o tópicos dentro de un conjunto de datos textuales, llevaría a construir datos cualitativos expresados numéricamente.

Los datos cualitativos de naturaleza textual son los que pueden realizar un aporte más novedoso y conceptualmente importante en el estudio de la diferenciación educativa. Nos referimos a los datos cualitativos registrados inicialmente en forma textual, o transformados en textos, bien se mantenga su naturaleza textual a lo largo del proceso de análisis o bien se transformen posteriormente en valores numéricos. Son precisamente este tipo de datos los que favorecen la explicación contextualizada de las diferencias, los que permiten profundizar en los elementos singulares de la diferenciación educativa, o los que aportan una visión más global y/o integrada de este problema. Los datos cualitativos registrados directamente mediante símbolos numéricos no realizan una aportación diferente a la que representan los datos cuantitativos recogidos mediante tests, inventarios, escalas, cuestionarios, etc.

ANÁLISIS DE DATOS CUALITATIVOS

Los datos que obtiene el investigador no suelen ofrecer suficiente información para resolver los complejos problemas planteados en la investigación sobre la diferenciación educativa, si no son organizados y manipulados en alguna forma. La información que aportan los datos *per se* no permite al investigador establecer relaciones, interpretar y extraer significados relevantes de cara al problema abordado, si no se realizan determinadas actividades y operaciones. En este sentido amplio solemos hablar de análisis de datos.

Las ideas básicas presentes en el concepto de análisis sostenido por una mayoría de autores hacen referencia al examen de las partes de algo por separado y de las relaciones existentes entre ellas, para reconstruir un todo significativo (Spradley, 1980): El análisis de datos es una actividad que implica un *conjunto de manipulaciones, transformaciones, reflexiones y comprobaciones realizadas a partir de los datos con el fin de extraer significado relevante para un problema de investigación*. Según Miles y Huberman (1984b), todas estas acciones presentes en un proceso de análisis de datos pueden resumirse en tres tareas básicas estrechamente relacionadas: reducción de la información, disposición o presentación de la misma, obtención de resultados y verificación de conclusiones (ver Figura 1):

Tradicionalmente se suele caracterizar el proceso de investigación por su linealidad temporal. Las etapas que lo componen tienen un sentido secuencial en el que el análisis de datos constituye una fase que sigue a la recogida de los datos y precede a su interpretación. Sin embargo, a menudo, en las notas de campo y junto a la descripción de situaciones educativas suelen aparecer juicios, opiniones, sospechas, dudas, reflexiones e interpretaciones del investigador, haciendo que el análisis y la elaboración de los datos se alternen o superpongan en el proceso de investigación. Lofland y Lofland (1984:132) presentan de un modo gráfico el desarrollo simultáneo de análisis y recogida de datos (ver Figura 2): Para estos autores, la

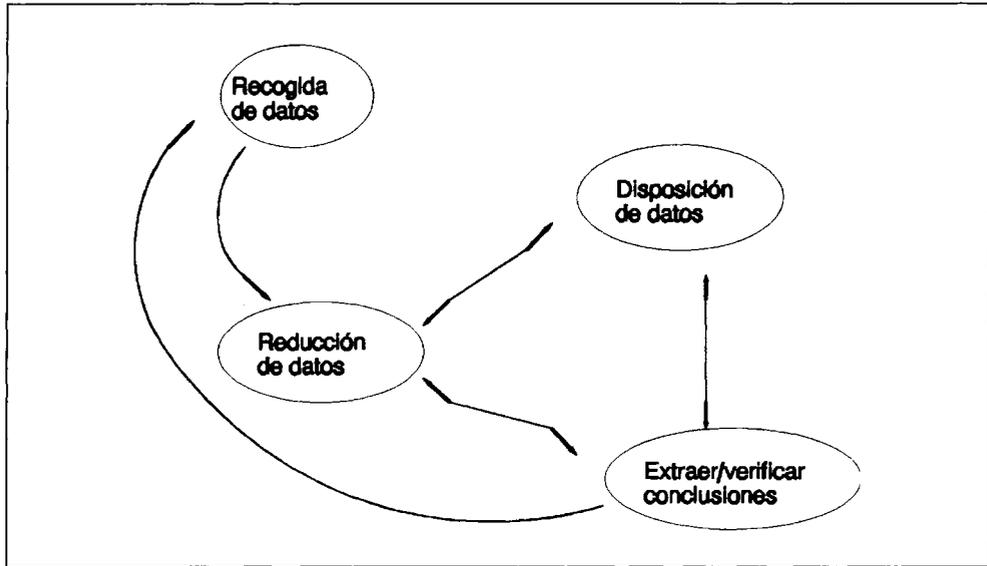


Figura 1
Componentes del análisis de datos (Miles y Huberman, 1984b: 23).

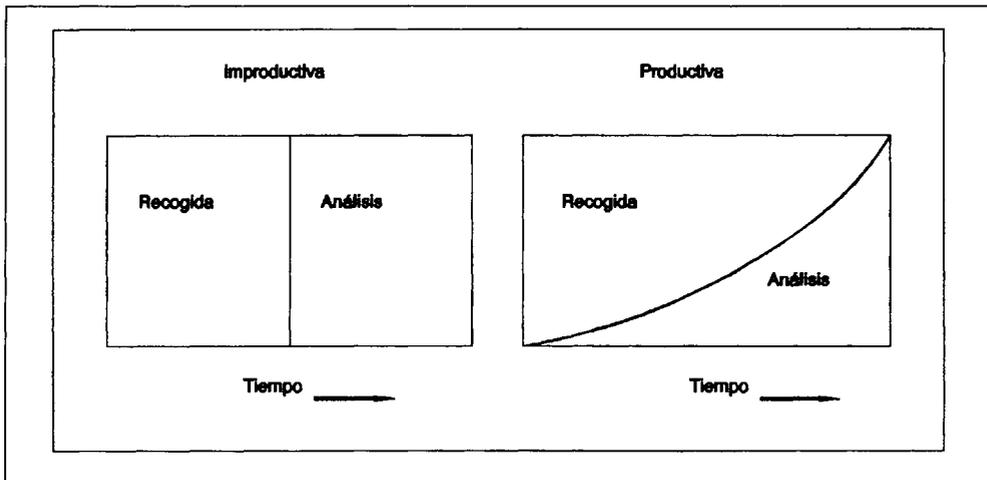


Figura 2
Relaciones productiva e improductiva entre recogida y análisis de datos (Lofland y Lofland, 1984: 132).

consideración de ambos procesos como fases separadas del trabajo del investigador resulta poco productiva, inclinándose por la concomitancia de tales tareas con un progresivo incremento del análisis, que prosigue una vez finalizada la recogida de

datos. Durante este último período se desarrollan y ponen en orden ideas surgidas previamente.

El bajo nivel de elaboración y la riqueza de significados intrínsecos a los datos cualitativos de naturaleza textual ha hecho de su análisis una tarea que ofrece no pocas dificultades. Enfrentarse al análisis de este tipo de datos ha representado para muchos investigadores una aproximación intuitiva hasta lograr estrategias particulares adecuadas a sus necesidades de análisis. La escasa definición de los métodos de análisis así generados hace que muchos investigadores vean a aquéllos simplemente como una «molestia atractiva» (Miles, 1985): La indefinición existente en relación a los procedimientos de análisis se ve agravada por el carácter plural de la investigación cualitativa, en la que coexisten un amplio número de tradiciones de investigación (Jacob, 1988).

Salvo quizás en el caso de los datos cualitativos textuales transformados en valores numéricos, no siempre el investigador cuenta con vías definidas o convenciones claras que orienten el modo de analizar los datos cualitativos, por lo que los procedimientos utilizados son de tal variedad y singularidad que a veces resultan particulares de un solo investigador o grupo de investigadores. Tanto es así que podría afirmarse que el único punto de acuerdo entre los investigadores es la idea de que el análisis es el proceso de extraer sentido de los datos (Tesch, 1990): Aún en el caso de que los investigadores lleguen a explicitar sus métodos, la falta de un lenguaje común hace que resulten poco claros. No sólo existe, por tanto, desacuerdo en el modo en que los investigadores realizan el análisis de sus datos cualitativos, sino también en la terminología que usan para referirse a las actividades y operaciones implicadas en él. Ello supone una dificultad añadida para el empleo de las técnicas de análisis por otros investigadores.

En síntesis, frente al carácter sistemático y formalizado de las técnicas de análisis aplicadas a los datos cuantitativos, los métodos usados con datos cualitativos presentan una gran diversidad y se basan con frecuencia en *la intuición y la experiencia del investigador*. A veces se ha afirmado que la tarea del analista es más un arte que una técnica, lo que supone sumar a las habilidades y tareas propias de todo proceso analítico otras de carácter especial. Para Woods (1989), el analista tendría que poseer actitudes y cualidades de *creatividad, imaginación* y un «*olfato de detective*» para extraer temas de interés a partir de indicios observados y llegar a producciones teóricas. Glaser y Strauss (1967) añaden a la imaginación la *ingenuidad*, y hablan de la *sensibilidad teórica* como característica que ha de poseer el investigador para conceptualizar y formular una teoría que emerja de los datos cualitativos. Otros aluden a la habilidad del investigador para *trascender su propia perspectiva* y llegar a conocer las perspectivas de aquellos a quienes está estudiando (Wilson, 1987): Autoras como Goetz y Lecompte (1988) hablan de la necesidad de contar con un *pensamiento divergente* y un *talento artístico*, aunque no refiriéndose a todo el proceso de análisis sino a algunos de los elementos instrumentales presentes en la interpretación (metáforas, analogías, etc.).

La singularidad y privacidad de los métodos de análisis cualitativos y la falta de

detalle con que son descritos en los informes de investigación, concentran parte de las críticas relativas a la credibilidad en la investigación cualitativa. Diversos autores han intentado dar rigor y exhaustividad a las técnicas cualitativas de análisis, tratando de que se aproximen en sistematización y fiabilidad a los procedimientos cuantitativos, si bien, como afirman Goetz y Lecompte (1988), hay autores que rechazan ese rigor y sistematización procesual por considerar que esclerotizan todo el proceso en perjuicio de su naturaleza intuitiva y creativa. Estos intentos de sistematización, en el análisis de datos cualitativos, se incluyen en lo que Firestone y Dawson (1982) denominan enfoques «de procedimiento», que incluyen una serie de reglas y pasos a seguir, frente a enfoques «intuitivos» o «intersubjetivos», basados respectivamente en la intuición y en la interacción entre investigadores o entre investigadores y participantes en los contextos estudiados. Los enfoques de procedimiento hacen esfuerzos por describir, con mayor o menor detalle, una serie de operaciones y pasos concretos, ofrecen consejos, recomendaciones y advertencias para el manejo, la disposición o la presentación de los datos. El avance que representan tales aportaciones en el análisis de datos cualitativos no ha evitado sin embargo que aún sean muchos los investigadores que recurren en algún momento de su proceso analítico a las técnicas estadísticas, con la intención de complementar o contrastar las conclusiones obtenidas por otras vías.

Efectivamente, la razón que apoya el uso de técnicas estadísticas en el análisis de datos cualitativos, de suerte que se sustituyan las palabras por números, puede encontrarse en el hecho de que aunque las palabras aportan mayor significado, los números son procesados con mayor economía. Por esta razón, la mayor parte de los investigadores prefiere trabajar con números, trasladando los textos a números lo más rápidamente posible. Para los investigadores que toman este último camino, el terreno está bien marcado por teorías de medición, reglas de decisión, niveles de confianza, algoritmos para el cálculo, etc.

De lo dicho anteriormente se deduce la existencia de dos tipos de datos cualitativos: los datos que conservan su naturaleza textual y los datos textuales que son transformados en valores numéricos. También han de considerarse como datos cualitativos, aquellos que se expresan directamente en forma numérica y están medidos en una escala nominal. No obstante, como ya se ha apuntado con anterioridad, este último tipo de datos cualitativos no representa una aportación novedosa o informativamente diferente a la que desde hace tiempo vienen ofreciendo los datos cuantitativos en el estudio de la diferenciación educativa.

La utilización que el investigador educativo hace de los datos cualitativos, manteniendo en éstos su naturaleza textual o bien transformándolos en valores numéricos, ha generado en la actualidad dos enfoques de análisis diferentes. Así, el análisis de datos cualitativos de naturaleza textual supone un proceso por el que se reduce y dispone la información recogida, se obtienen resultados y verifican conclusiones acudiendo a procedimientos preferentemente no estadísticos. Por su parte, el análisis de datos cualitativos transformados en valores numéricos se realiza mediante procedimientos estadísticos. A pesar de esta diferenciación, para el análisis de datos

cualitativos, es posible reconocer la existencia de un proceso común, en muchos de sus elementos, tanto a los datos de naturaleza textual como a los transformados en valores numéricos.

PROCESO GENERAL DE ANÁLISIS DE DATOS CUALITATIVOS

El análisis de datos cualitativos va a ser considerado aquí como un proceso realizado con un cierto grado de sistematización que, a veces, permanece implícita en las actuaciones emprendidas por el investigador. En este sentido, resulta difícil

CUADRO I
PROCESO GENERAL DE ANÁLISIS DE DATOS CUALITATIVOS

TAREAS	ACTIVIDADES	OPERACIONES
	Separación de unidades	Determinación de criterios de separación físicos, temáticos, gramaticales, conversacionales y sociales.
Reducción de datos	Identificación y clasificación de elementos	Categorización y codificación
	Síntesis y agrupamiento	Agrupamiento físico, creación de metacategorías, obtención de estadísticos, métodos estadísticos de agrupamiento y síntesis.
Disposición y transformación de datos	Disposición	Elaboración de tablas numéricas, gráficos, modelos, matrices y sistemas de redes
	Transformación	Expresión de los datos en otro lenguaje (numérico, gráfico)
Obtención de resultados y verificación de conclusiones	Proceso para obtener resultados	<i>Datos textuales:</i> descripción e interpretación; recuento y coocurrencia de códigos; comparación y contextualización. <i>Datos transformados en valores numéricos:</i> técnicas estadísticas; comparación y contextualización
	Proceso para alcanzar conclusiones	<i>Datos textuales:</i> consolidación teórica, aplicación de otras teorías, uso de metáforas y analogías, síntesis con resultados de otros investigadores. <i>Datos transformados en valores numéricos:</i> uso de reglas de decisión (comparación de los resultados con modelos teóricos; recurso a la perspicacia y experiencia del analista).
	Verificación de conclusiones	Comprobación o incremento de la validez mediante presencia prolongada en el campo, intercambio de opiniones con otros investigadores, triangulación, comprobación con los participantes, establecimiento de adecuación referencial, ponderación de la evidencia, comprobación de la coherencia estructural.

hablar de una estrategia o procedimiento general de análisis de datos cualitativos, fuera de lo que pueda inferirse a partir de las acciones identificadas en un análisis ya realizado. No obstante, tomando como base estas inferencias, es posible establecer una serie de tareas u operaciones que constituyen el proceso analítico básico, común a la mayoría de los estudios en que se trabaja con datos cualitativos. Estas tareas serían: reducción, disposición y transformación, y obtención de resultados y verificación de conclusiones.

En cada una de estas tareas es posible distinguir, asimismo, una serie de actividades y operaciones concretas que son realizadas durante el análisis de datos, aunque no necesariamente todas ellas estén presentes en el trabajo de cada analista. En ocasiones, determinadas actividades pueden extenderse hasta constituir por sí mismas el proceso de análisis o, por el contrario, pueden no ser tenidas en cuenta en el tratamiento de los datos, de acuerdo con los objetivos del trabajo, el enfoque de investigación, las características del investigador... etc. Entre ellas no siempre se establece una sucesión en el tiempo, y pueden ocurrir de forma simultánea, o incluso estar presentes varias de ellas dentro de un mismo tipo de tarea. El marco general desde el que suelen analizarse los datos cualitativos podría esquematizarse en las siguientes tareas, actividades y operaciones recogidas en el cuadro I.

A) Reducción de datos

La limitada capacidad humana para procesar grandes cantidades de información hace necesaria su reducción a unidades elementales, fácilmente analizables, pero éstas a su vez, han de ser comprensivas, relevantes y significativamente densas. Para ello se suele recurrir a tres actividades básicas: separación, clasificación y síntesis, que incluyen operaciones tales como la categorización y codificación. Estas tres actividades son complementarias y, a veces, indisociables en la práctica analítica del investigador.

La reducción de datos no es una tarea asociada necesariamente a una fase inicial del análisis. En momentos avanzados del análisis, la reducción de datos puede hacerse sobre los resultados de las operaciones analíticas realizadas con anterioridad tomando como base los datos originales. Es posible advertir la función reductora incluso antes de contar con los datos. Miles y Huberman (1984b) consideran que puede darse una reducción de datos anticipada cuando focalizamos y delimitamos la recogida de datos, aspectos constituyentes de una forma de preanálisis que comporta descartar determinadas variables y relaciones y atender a otras. De alguna manera, los datos recogidos suponen un conjunto conscientemente reducido del universo de datos que podrían haberse obtenido. Finalmente, la reducción está presente cuando el investigador resume o esquematiza sus notas de campo o cuando selecciona entre los datos obtenidos aquellos que considera relevantes para sus cuestiones de investigación.

A.1. Separación de unidades

De acuerdo con el significado básico del término «análisis», una de sus primeras actividades es la separación de las unidades que conforman el cuerpo de datos objeto de estudio. Tesch (1990:95) afirma, como uno de los principios y prácticas más características del análisis cualitativo, que los datos pueden ser segmentados, es decir, «divididos en unidades relevantes y significativas».

Los investigadores que trabajan con datos cualitativos, en ocasiones, no suelen explicitar en sus estudios los criterios seguidos para la división del corpus de datos en unidades, e incluso se llega a obviar este importante aspecto. Lofland y Lofland (1984), por ejemplo, describen minuciosamente la operación de cortar los textos en fragmentos (unidades) pero omiten el modo en que llegan a ser delimitados tales fragmentos. No obstante, esta separación puede realizarse sobre la base de criterios físicos (espaciales y temporales), temáticos, gramaticales, conversacionales y sociales.

Las consideraciones *espaciales o temporales* permiten definir unidades de carácter artificial. De acuerdo con tales criterios, constituirían unidades los bloques de un determinado número de líneas, las páginas, los bloques de minutos, etc. Es más frecuente, y probablemente más adecuado, considerar unidades en función del *tema abordado*. Erickson (1982), al analizar datos audiovisuales, propone un modelo de análisis en el que se desciende desde el todo a las partes: partiendo del suceso como un todo, diferencia elementos constituyentes cada vez más pequeños, considerando conversaciones, sucesos, actividades que ocurren en la situación observada; a partir de ellos, es posible encontrar segmentos que hablan de un mismo tema. Sin embargo, el criterio temático conlleva una pequeña dificultad cuando las unidades son objeto de cómputo para posteriores análisis. Podríamos encontrarnos tanto fragmentos de gran extensión como fragmentos reducidos que hablen de un mismo tema, resultando unidades que llegan a ser consideradas equivalentes en análisis ulteriores, a pesar de que no lo son. Este problema ha sido soslayado en el análisis de contenido mediante la distinción entre unidades de registro y unidades de enumeración (Cartwright, 1978): Mientras las primeras determinan fragmentos referidos a un mismo tópico, las segundas aluden a la unidad sobre la base de la cual se realiza el recuento.

Las unidades pueden establecerse a partir de *criterios gramaticales* (el párrafo, la oración o la palabra podrían ser considerados las unidades básicas de un texto) o *conversacionales* (declaraciones, turnos de palabra, etc.): En ellas no es necesario un juicio sobre el significado para establecer los límites de la unidad. En algunas investigaciones, las unidades también se han elaborado a partir de *criterios sociales*. Cada unidad se haría corresponder con el segmento de texto generado por la intervención de un sujeto con determinado status o rol social (p.e. profesor, director, alumno, inspector, etc.):

A.2. Identificación y clasificación

La identificación y clasificación de elementos consiste en examinar las unidades de datos para encontrar en ellas determinados componentes generalmente temáticos que nos permitan clasificarlas en una u otra categoría de contenido. Las operaciones conocidas como categorización y codificación son las más representativas dentro de las actividades de identificación y clasificación.

La *categorización* hace posible agrupar conceptualmente las unidades que son cubiertas por un mismo tópico. Mucchielli (1988:34) define la categoría como «una noción general que representa a un conjunto o una clase de significados». Para Bogdan y Biklen (1982), las categorías pueden referirse a situaciones y contextos, actividades y acontecimientos, relaciones entre personas, comportamientos, perspectivas sobre un problema, métodos y estrategias, procesos. La categorización se realiza conjuntamente a la división en unidades cuando ésta se hace atendiendo a criterios temáticos. Si una unidad es separada por referirse a determinado tópico, automáticamente puede ser incluida en la categoría correspondiente a ese tópico.

El proceso de *codificación* no es más que la operación concreta por la que se asigna a cada unidad un indicativo (código) propio de la categoría en la que se incluye. Es el proceso físico, manipulativo, mediante el cual se realiza la categorización. Por esta razón, no es extraño que el estudio de ambas operaciones se aborde a menudo de forma conjunta, y que incluso en algún momento lleguen a identificarse como un único proceso, aspecto que queda reflejado en afirmaciones tales como «categorizar es asignar etiquetas a algo» (Bliss, Monk y Ogborn, 1983:8), o «los códigos son categorías» (Miles y Huberman, 1984b:56).

Los códigos, que representan a las categorías, pueden ser números con los que se marca cada unidad de datos (Bogdan y Biklen, 1982), aunque es más frecuente utilizar abreviaturas de palabras como categorías (Tesch, 1990): Por tanto, físicamente el código es «una abreviatura o símbolo aplicado a un segmento de texto — muchas veces una frase o un párrafo de notas de campo transcritas— con el fin de clasificarlo» (Miles y Huberman, 1984b:56).

Las categorías utilizadas en un estudio pueden establecerse a priori, de acuerdo con un marco teórico y conceptual previo, las cuestiones o hipótesis que guían la investigación, las categorías ya usadas en estudios de otros investigadores o los instrumentos de investigación empleados (el uso de cuestionarios o guiones de entrevista, por ejemplo, permite utilizar las cuestiones planteadas como fuente de categorías o modalidades útiles para organizar los datos): En ocasiones, las categorías que utiliza un investigador se relacionan entre sí dando lugar a estructuras de cierta complejidad conceptual conocidas como *sistemas de categorías*. Estos sistemas suponen un nivel explicativo superior al de la simple enumeración de temas o categorías. La realidad educativa analizada a través de esos sistemas de categorías se explica desde la localización física, el orden, la covariación, la causalidad, etc.

El partir de un sistema de categorías a priori no es una condición suficiente, sino un requisito previo para conseguir una misma codificación en distintos codificado-

res. Es preciso además una definición operativa de categorías, ofreciendo reglas que especifiquen los aspectos del contenido que deben tomarse como criterio de pertenencia a una categoría (Cartwright, 1978): Ha habido intentos de proporcionar sistemas de categorías válidos para múltiples estudios, como el ampliamente difundido sistema de categorías de Bales para el análisis de conductas lingüísticas en un grupo (ver Anguera, 1978): Apoyadas en este planteamiento a priori se han diseñado aplicaciones informáticas para el análisis de datos textuales capaces de realizar el proceso de codificación de modo automático, aunque con la limitación de que las categorías deben estar definidas por la aparición de un término o frase en los datos que señalen la relevancia del fragmento para una categoría (Tesch, 1990): La codificación automática eliminaría los efectos asociados a las características personales y circunstanciales del codificador, las cuales afectan al proceso de reducción de datos (Molly, 1984).

Los sistemas de categorías también pueden elaborarse inductivamente a partir de los propios datos. Al examinar los datos nos preguntaríamos por el tópico capaz de cubrir cada unidad, generando así una categoría. Posteriores comparaciones entre los datos agrupados bajo un mismo tópico o entre tópicos permitirían refinar las categorías emergentes (Glaser y Strauss, 1967): En el proceso denominado «codificación abierta» (Strauss, 1987), se parte de la búsqueda de conceptos que traten de cubrir los datos. En muchos casos, el sistema de categorías deja de ser únicamente una herramienta que utiliza el investigador para organizar sus datos y se constituye por sí mismo en un producto del análisis, un resultado de investigación. Para Spradley (1979, 1980), uno de los objetivos es llegar a elicitar las categorías (dominios, en la terminología de este autor) en que se clasifican los términos de una cultura y reducir de ese modo la complejidad de la experiencia humana. Su interés está en descubrir cómo las personas organizan su conocimiento, por lo que las categorías son un fin en sí mismas; no tiene sentido imponer categorías analíticas a los datos obtenidos durante el estudio.

Por lo general, procedimientos deductivos e inductivos se combinan en el trabajo del analista, de modo que a partir de unas categorías a priori se producen modificaciones para adaptarlas al conjunto de datos a que están siendo aplicadas. Tanto si las categorías son definidas a partir de los datos como si son preelaboradas (en función del estudio o tomados de sistemas generales), codificación y categorización no se desarrollan linealmente, sino que son procesos recurrentes, durante los cuales el sistema de categorías puede ser alterado. A medida que el investigador posee un mayor conocimiento del material que analiza, puede alterarse de forma no deliberada el modo en que se asignan los códigos. Cuando surjan incongruencias derivadas de la aplicación inconsistente de códigos, habrá que revisar lo realizado hasta el momento. De hecho, durante la codificación es posible advertir que determinados códigos podrían haber sido agrupados en uno solo o, por el contrario, que cabría dividir uno de ellos en otros diversos, ciertas categorías podrían «necesitar ser renombradas, modificadas en su contenido o sustituidas por otras nuevas» (Tesch, 1990:91):

Miles y Huberman (1984b) señalan que la codificación puede llevarse a cabo en diferentes momentos de la investigación y distinguen entre códigos *descriptivos* que atribuyen una unidad a una clase de fenómenos y son utilizados en un primer momento de la codificación, y códigos con un mayor contenido inferencial (*interpretativos y explicativos*) utilizados posteriormente. Por otra parte, los sistemas de categorías no son únicos, sino que mediante las categorías no buscamos sino extraer distinciones y nombrarlas, y esas distinciones pueden hacerse en varias dimensiones independientes (Bliss, Monk y Ogborn, 1983): Según el aspecto en que nos fijemos, un texto puede ser categorizado de distinto modo.

Autores que ven en la categorización un modo de transformar los datos textuales en datos susceptibles de medición y tratamiento estadístico (Ander-Egg, 1990; Cartwright, 1978; Kerlinger, 1975; Sánchez Carrión, 1985, etc.) demandan de los sistemas de categorías una corrección lógica que se alcanza con los requerimientos clásicos de *exhaustividad* de las categorías (cualquier unidad puede ser ubicada en alguna de las categorías), *exclusión mutua* (cada unidad se incluye en una sola categoría) o el estar elaboradas desde un *único principio clasificatorio*. Mucchielli (1988) habla también de la *objetividad* (inteligibles para distintos codificadores), y la *pertinencia* en relación a los objetivos del estudio y al propio contenido analizado. Sin embargo, concepciones más recientes en la investigación en ciencias humanas abogan por categorías que no están claramente delimitadas unas de otras, con superposición entre las áreas temáticas que cubren. Se habla en estos casos de *membreía* en una categoría, para expresar en qué grado determinada unidad puede formar parte o ser miembro de ella, y de la noción de *prototipo*, que representa el tipo ideal con un cien por ciento de pertenencia a una sola categoría. Tal elemento no existe realmente; es un constructo mental que sirve como punto de referencia al que comparar las unidades de datos reales, para determinar su grado de membreía (Tesch, 1990).

A.3. Síntesis y agrupamiento

La identificación y clasificación de elementos están estrechamente unidas a la síntesis. Cuando categorizamos estamos sintetizando diferentes unidades de datos en un mismo tópico o concepto teórico. La categorización supone en sí misma una operación conceptual de síntesis, en la que se reducen diversas unidades a un concepto que las representa. En el caso de los datos cualitativos textuales, la síntesis permite realizar un agrupamiento físico en el que se reúnen todas las unidades que forman parte de la misma categoría. En el caso de los datos cualitativos transformados en valores numéricos, la unión de modalidades de variables supone una operación de agrupamiento, mientras la obtención de estadísticos de tendencia central supone procesos de síntesis.

En el caso del análisis de contenido automático se ha recurrido a dos procedimientos para llegar a agrupar unidades (generalmente palabras): Los *sistemas a priori* buscan palabras identificadas con conceptos teóricos relevantes para el pro-

blema investigado (Mochmann, 1985): Es necesario definir previamente la relación entre conceptos y palabras, creando lo que se denomina un diccionario de análisis de contenido. Mediante este diccionario el investigador sintetiza previamente el significado de las palabras elaborando un concepto que las agrupa. El ordenador se encargará de localizar todas las palabras que se agrupan bajo cada concepto y ofrecer sus frecuencias de aparición en las unidades consideradas para permitir cálculos estadísticos. Existen diccionarios generales, apoyados en teorías de gran alcance, y también diccionarios especializados o incluso específicos de cada estudio (Mohler, 1985):

Otros enfoques, denominados *empíricos*, agrupan automáticamente las palabras (excluyendo las palabras funcionales: artículos, conjunciones, etc.) sin utilizar categorías preestablecidas, y evitan así toda influencia del investigador en el agrupamiento resultante. Las síntesis (categorías) surgen mediante procedimientos estadísticos como los análisis factoriales o los análisis de conglomerados realizados sobre matrices de correlaciones obtenidas a partir de la frecuencia que cada palabra alcanza en una unidad diferenciada en el texto (Iker y Harway, 1969):

Los procesos de agrupamiento y síntesis también están presentes en el proceso de análisis cuando, por ejemplo, agrupamos varios códigos en un metacódigo o cuando sintetizamos la información contenida en varias categorías que tienen algo en común en una metacategoría, o elaboramos un sistema de indización.

Las actividades y operaciones revisadas hasta aquí constituyen modos de contribuir a la reducción de datos que en el fondo presupone todo análisis: se parte de un conjunto amplio y complejo de datos para llegar a unidades elementales más manejables que permitan establecer relaciones y sacar conclusiones.

B) Disposición y transformación de los datos

Como afirma Levine (1985), al aportar razones para apoyar la conveniencia de un análisis sistemático de la información verbal, los datos de este tipo no son recogidos siguiendo, en una secuencia lineal, a los problemas de la investigación, sino en función de la interacción del investigador y el informante o el contexto de estudio. Además, los textos narrativos resultan dispersos, ofrecen una visión secuencial de hechos simultáneos, dificultan el examen de varias variables a la vez y están vagamente ordenados (Miles y Huberman, 1984b): Cuando los datos cualitativos se transforman numéricamente el analista también se encuentra ante dificultades para estudiar directamente dichos datos, de modo que sea posible describir la tendencia central de cada variable o para identificar directamente la intensidad de las posibles relaciones entre las mismas.

Dadas estas dificultades, es necesaria una *transformación y disposición* de los datos que permita presentarlos de manera abarcable y operativa de cara a resolver las cuestiones de investigación. Disposición y transformación son dos actividades que facilitan el examen y la comprensión de los datos, permitiendo extraer conclusiones, y condicionando posteriores decisiones en el proceso de análisis. La dispo-

sición supone organizar los datos, presentándolos en alguna forma espacial ordenada, de tal modo que se simplifique la información y se posibilite su procesamiento posterior. Existen diferentes procedimientos aplicables a la disposición de los datos cualitativos.

Un primer procedimiento lo constituyen las *tablas numéricas*, entre las que podemos diferenciar los siguientes tipos (Lebart, Morineau y Fenelon, 1985):

- a) Tablas de contingencia, en las que una misma población puede ser objeto de dos particiones, que se sitúan en filas y columnas respectivamente. En la intersección de éstas aparecerá el número de sujetos o casos que pertenece simultáneamente a dos partes consideradas. Estas tablas tienen por finalidad estudiar la relación entre dos o más variables de tipo nominal u ordinal.
- b) Tablas de incidencia, o también denominadas tablas de presencia-ausencia. Sujetos y variables son trasladados a filas y columnas. Según un individuo u objeto presente o no determinada característica o atributo, se consignará 1 ó 0 en la celda de intersección correspondiente.
- c) Tablas de valores numéricos, recogiendo en las intersecciones el valor de las variables ubicadas en las columnas para los individuos colocados en las filas.

La presentación de datos en forma numérica puede también darse una vez que se ha llevado a cabo un cierto procesamiento estadístico de la información, recurriendo en estos casos a tablas que recogen distribuciones de frecuencias, proporciones, porcentajes, etc. Otras tablas o matrices numéricas presentan ordenadamente datos tales como correlaciones, pesos y puntuaciones factoriales, distancias entre sujetos, etc. obtenidos tras diversos análisis. En cualquier caso, las tablas cumplen una función analítica. Su finalidad no es almacenar datos, sino comunicarlos en algún modo ordenado que permita extraer conclusiones (Wainer, 1992):

Un segundo procedimiento de disposición de los datos lo constituyen los *gráficos*, medios para presentar información que ha podido ser objeto de un mayor o menor grado de elaboración. Los histogramas, curvas de nivel, diagramas de sectores circulares son utilizados para describir gráficamente distribuciones de frecuencias. Gráficos de caja, postes telegráficos, tallo y hojas, matrices ordenables, matrices ponderadas, etc. son procedimientos que se utilizan en la disposición de los datos cualitativos. Asimismo, técnicas como el análisis factorial de correspondencias, el escalamiento multidimensional o el análisis de conglomerados conllevan representaciones gráficas que muestran la situación de los elementos respecto a ejes factoriales, su posicionamiento en un espacio geométrico, o el modo en que se agrupan (dendograma): Los gráficos permiten no sólo presentar los datos, sino advertir relaciones y descubrir su estructura profunda (Wainer, 1992).

Como un tercer grupo de procedimientos, pueden destacarse las técnicas de *modelado gráfico* muy extendidas entre los investigadores que trabajan con datos cualitativos de naturaleza verbal. Spradley (1979,1980) utiliza representaciones taxonómicas en forma de perfil general («outline»), diagrama de caja («box diagram»)

o diagrama en árbol («lines and nodes»): También se apoya en disposiciones denominadas paradigmas, que se acercan más a un formato de matriz, en las que se señala la presencia o no de ciertos atributos (columnas) en un grupo de términos considerados (filas): Otra técnica de modelado gráfico la constituye los diagramas esquemáticos, «representaciones gráficas o imágenes visuales de las relaciones entre conceptos» (Strauss y Corbin, 1990:198).

Las *matrices* constituyen otra forma de disposición gráfica de los datos. Consisten en tablas de doble entrada en cuyas celdas se aloja una breve información verbal (fragmentos de texto, citas, frases, abreviaciones o figuras simbólicas), de acuerdo con los aspectos especificados por filas y columnas. Es frecuente construirlas para sintetizar los fragmentos codificados en una misma categoría (columna) para diferentes sujetos, situaciones, casos, etc. (filas): Algunas de las dimensiones sobre la base de las cuales se elaboran pueden implicar la ordenación de fenómenos o componentes de fenómenos a lo largo del tiempo, en función de los roles desempeñados, etc. Determinadas matrices presentan procesos y productos, relaciones de causa-efecto, agrupamientos conceptuales o empíricos de datos.

Bliss, Monk y Ogborn (1983) proponen un *sistema de redes* («networks») mediante las cuales tratan de clarificar y presentar las distinciones que delimitan a las diversas categorías consideradas en el análisis del material cualitativo, las cuales a su vez pueden estar subdivididas. Para ello usan una notación precisa, con la que indican nombres de categorías, subdivisiones de éstas, diferencias, circunstancias, restricciones, etc.

La presentación de un mismo corpus de datos puede hacerse en múltiples formas, según el objetivo que el analista persiga con ello. A veces, el interés es exclusivamente descriptivo, tratando de ofrecer la información disponible y permitir una visión de conjunto que lleve a análisis o presentaciones más detalladas. En otras ocasiones, ese interés se centra en la interpretación y el analista trata de presentar los datos de forma que permitan explicar y crear generalizaciones, desarrollar teorías, etc. También la disposición de los datos puede favorecer la verificación de supuestos, teorías, etc. o el estudio de prácticas e intervenciones educativas.

Cuando la disposición de los datos no sólo implica un cambio en la forma de presentarlos espacialmente sino también una modificación del lenguaje, hablamos de transformación de los datos. Las transformaciones más frecuentes son las que permiten pasar de una expresión verbal de los datos a una expresión numérica, o bien de cualquiera de estas formas de expresión a una de tipo gráfico. Se realizan transformaciones cuando contamos códigos, palabras, segmentos de palabras,... y proseguimos el análisis a partir del recuento efectuado y no de la forma verbal en que aparecían los datos. También son ejemplos de transformación aquellos procedimientos de disposición que implican una expresión gráfica de los datos (histogramas, modelos gráficos, representaciones taxonómicas): Como también ocurre con los procedimientos empleados para la disposición, la transformación de los datos facilita la descripción, interpretación, verificación o valoración de las diferencias educativas.

C) Obtención de resultados y verificación de conclusiones

C.1. *Proceso para obtener resultados*

Un importante grupo de actividades del proceso de análisis se dirige a la obtención de resultados, en relación a los objetivos particulares de cada investigación. Obtener resultados implicaría, de acuerdo con las connotaciones semánticas del término análisis, ensamblar de nuevo los elementos ya diferenciados en el proceso analítico para reconstruir un todo estructurado y significativo. Van Maanen (1985) denomina a los resultados «conceptos de segundo orden», pues se construyen a partir de los datos, o conceptos de primer orden. La obtención de resultados no tiene lugar necesariamente al término del proceso de análisis, sino que éstos son extraídos por el investigador durante todo el proceso de recogida de datos y durante otros momentos del análisis (Miles y Huberman, 1984b): Las tareas de reducción y presentación de datos le permiten hacer afirmaciones que progresivamente avanzan desde lo descriptivo a lo explicativo y desde lo concreto a lo abstracto. Las categorías obtenidas en el proceso de reducción pueden considerarse, asimismo, resultados del estudio y, como se señaló más arriba, la ordenación de los datos en una matriz, un cuadro, una figura, etc. presupone también algún tipo de conocimiento respecto a su estructura.

Cuando se analizan los datos cualitativos conservando su naturaleza textual, los resultados se obtienen en función del modo en que se dispone y organiza la información, a partir del recuento o de la coocurrencia de determinados códigos. Los resultados también surgen a partir de una interpretación de los datos una vez reducidos y dispuestos en alguna forma: el investigador alcanza una visión global de los datos que no necesariamente coincidirá con la de otros investigadores que se enfrentaran a los mismos datos. La lectura de una gran parte de los trabajos que se publican sugeriría que los resultados emergen directamente a partir del corpus de datos, sobre todo una vez que éste ha sido adecuadamente reducido y presentado.

Operaciones que lleva a cabo el investigador para obtener resultados a partir de datos de naturaleza textual son la comparación y la contextualización. La primera de ellas, permite destacar las semejanzas y diferencias entre las unidades incluidas en una categoría, o entre distintas categorías, y hace posible la formulación de sus propiedades fundamentales, a partir de las cuales puede llegarse a una definición, ilustración y verificación de esa categoría (Fielding y Fielding, 1986): La comparación facilita la obtención de resultados a partir de las matrices y otras figuras usadas para la presentación de los datos. La comparación de filas o columnas de una matriz, por ejemplo, constituye la principal vía para extraer resultados (Miles y Huberman, 1984b): Los resultados pueden emanar de la comparación con otros escenarios, casos, situaciones, etc. similares al estudiado. También la comparación puede realizarse respecto a un criterio particular (una meta u objetivo) o a una norma, que sirven de base para el establecimiento de un juicio de valor. Los resultados se obtienen además merced a la contextualización de los datos. Así, los

procesos, las relaciones, los sistemas, las opiniones presentes en los datos pueden ser referidos al marco cultural en el que se han generado, de suerte que lo que se describe, interpreta, verifica o evalúa forma parte de un todo más comprensivo que da unidad al conjunto de elementos considerados en el análisis.

Cuando se trata de datos cualitativos transformados en valores numéricos el proceso de obtención de resultados es más inequívoco y transparente. El investigador cuenta con un amplio repertorio de algoritmos de cálculo, técnicas estadísticas y operaciones matemáticas que traducen un razonamiento lógico de naturaleza formal, todos ellos le permiten obtener elaboraciones de segundo orden fácilmente replicables por otros investigadores. También con este tipo de datos se utiliza la comparación (inter o intra) como herramienta para obtener resultados. El análisis de tablas de contingencia, por ejemplo, se basa en la comparación del modo en que se distribuyen los datos por filas o por columnas. Por su parte, la contextualización está presente cuando disponemos espacialmente los datos poniendo de manifiesto la posición relativa de cada elemento respecto al conjunto.

C.2. *Proceso para alcanzar conclusiones*

Otro momento en el análisis de los datos supone la interpretación de los resultados obtenidos hasta *alcanzar conclusiones*. La interpretación nos lleva a conclusiones que representan una lectura más comprensiva de los resultados y que pueden conducir a la creación y explicación de generalizaciones, el planteamiento de nuevas cuestiones o problemas, el desarrollo de nuevos conceptos, la redefinición de conceptos ya existentes, la clarificación y comprensión de la complejidad de algo o la generación de una nueva teoría o explicación.

En el caso de los procedimientos de análisis de datos cualitativos que conservan la forma textual, pueden ser utilizadas estrategias como las propuestas por Goetz y Lecompte (1988) para alcanzar conclusiones a partir de la interpretación de los resultados. Estas estrategias, que suponen integrar los resultados en marcos teóricos y de investigación más amplios, serían las siguientes:

- Consolidación teórica. Consiste en confrontar los datos obtenidos con los marcos teóricos desarrollados originalmente, modificando éstos para intentar que se ajusten más a los datos. Este enfoque ha sido ampliamente desarrollado por Glaser y Strauss (1967), quienes han defendido la utilidad de generar teoría fundamentada en los datos.
- Aplicación de otras teorías. La búsqueda de otros marcos más generales en los que integrar los datos permite la generalización de los resultados.
- Uso de metáforas y analogías. Son medios útiles para establecer conexiones entre temas aparentemente no relacionados, o conectar los resultados a la teoría (Miles y Huberman, 1984b): Preguntarse «¿a qué me recuerda esto?», acerca de los diferentes aspectos estudiados puede ser una estrategia adecuada para encontrar analogías y metáforas (Bogdan y Biklen, 1982:153).
- Síntesis de los resultados con los obtenidos por otros investigadores, para lo

que es necesario comparar, contrastar, agregar, ordenar, establecer relaciones y especular. Es decir, se repite el proceso analítico pero esta vez con los resultados alcanzados y, por tanto, a niveles más elevados de abstracción y generalización.

En el análisis de datos cualitativos numéricos, el proceso que nos permite llegar a conclusiones a partir de los resultados obtenidos se ve facilitado por reglas de decisión previamente establecidas, basadas en la comparación de los resultados con modelos teóricos de distribución del estadístico de contraste. Los procedimientos estadísticos que no se basan en la inferencia también utilizan reglas de decisión, si bien ésta se apoya en parte en la experiencia y perspicacia del analista. Por ejemplo, ante la representación gráfica resultante de un análisis factorial de correspondencias, el analista tendría que ser capaz de identificar un rasgo o característica que es poseída con distinto signo y peso por los elementos situados a uno u otro lado de un determinado eje; el factor representado por ese eje puede quedar identificado con la característica en cuestión.

C.3. Verificación de conclusiones

Una vez alcanzadas las conclusiones de un estudio, es necesario verificarlas. Esta tarea supone comprobar el valor de verdad de los descubrimientos realizados, es decir, contrastar su validez. Entendida como un estado ideal que se pretende alcanzar y no como un logro, la validez ha sido definida desde distintas posiciones epistemológicas que han elaborado diferentes patrones para verificar las conclusiones de un estudio. Desde determinada posición epistemológica, identificable con programas de investigación positivistas o neopositivistas, se entiende que las conclusiones son válidas en la medida en que existe una realidad educativa que las sostiene o que al menos no las contradice. Para favorecer o incrementar la obtención de conclusiones válidas, desde esta posición, se hacen esfuerzos por conseguir que el proceso de investigación no se vea afectado por diferentes fuentes de error y que los resultados puedan ser generalizados.

Desde otra posición epistemológica, identificable con programas de investigación interpretativos, se considera que la validez es un ideal que se verifica en función de la concordancia entre las conclusiones extraídas durante el análisis y las perspectivas mantenidas por los sujetos o grupos estudiados. Podemos hablar de validez cuando las personas comparten —quizás tras un proceso de diálogo y negociación— una descripción o interpretación (Smith y Heshusius, 1986:9): Autores como Guba (1983) plantean además la necesidad de que se dé una adecuación entre datos e interpretaciones y entre proposiciones y referentes.

Cuando los datos cualitativos se analizan desde esta posición, la validez de las conclusiones puede incrementarse, o simplemente ser calibrada, con estrategias desarrolladas en las fases de diseño, recogida de datos, análisis de datos o redacción del informe (Dawson, 1982): Por el hecho de estar presente, el investigador provoca

en los participantes conductas que ordinariamente no habrían ocurrido y es, al mismo tiempo, influenciado por ellos. La *presencia prolongada en el campo* hace que el investigador forme parte de la situación estudiada permitiéndole soslayar los efectos de su presencia, le permite comprobar los prejuicios y percepciones propios o de los participantes, y comprender lo que es esencial o característico (Guba, 1983), haciendo posible el ajuste entre las categorías de análisis extraídas y la realidad. El *intercambio de opiniones* con otros investigadores, exponiéndose a sus preguntas y críticas puede ser útil para llegar a conclusiones válidas. La publicación de resultados constituye un modo de presentarlos a la revisión por parte de otros expertos (Goetz y Lecompte, 1988): La *triangulación* es un procedimiento ampliamente usado que consiste en el uso de diferentes fuentes de datos, investigadores, perspectivas o metodologías para contrastar los datos e interpretaciones (Denzin, 1978): Si diversas clases de datos conducen a la misma conclusión podemos estar algo más seguros de tal conclusión (Hammersley y Atkinson, 1983): Sin embargo, se han hecho objeciones a la triangulación en el caso en que métodos de diferentes perspectivas metodológicas son usados conjuntamente a pesar de responder a asunciones ontológicas y epistemológicas diferentes (Blaikie, 1991): La *comprobación con los participantes* supone contrastar los datos e interpretaciones de los mismos con los sujetos que constituyen la fuente de esos datos. Como profundos conocedores de la realidad que se investiga, los participantes podrían actuar como jueces que evalúen los principales descubrimientos de un estudio (Denzin, 1978): Pueden realizarse esas comprobaciones de los resultados frente a datos nuevos procedentes de otras fuentes (documentos, grabaciones, etc.) recogidos con tal fin durante el estudio, estrategia denominada por Guba (1983) *establecimiento de adecuación referencial*.

Miles y Huberman (1984b) proponen la *ponderación de la evidencia* como técnica que incrementaría la validez de los resultados. Su idea es la de considerar ciertos datos mejores que otros a la hora de extraer conclusiones, ya sea por su procedencia de determinados informantes, por las circunstancias en que fueron recogidos (al final o al principio de la estancia en el campo, revelados en privado o en presencia de otros informantes, etc.) o por haber sido validados mediante estrategias como la triangulación o la comprobación con los participantes. Debe ser comprobada la *coherencia estructural* para asegurarse de que entre los datos e interpretaciones no se dan contradicciones o incoherencias (Guba, 1983): Un modo de aproximarse a ella es la *búsqueda de evidencia negativa* (Miles y Huberman, 1984b), consistente en preguntarse, una vez alcanzadas las conclusiones, si hay datos que se oponen o son inconsistentes con esas conclusiones.

MÉTODOS DE ANÁLISIS DE DATOS CUALITATIVOS

La investigación sobre la diferenciación educativa puede apoyarse en diferentes métodos para el análisis de datos cualitativos. La elección de uno u otro procedimiento analítico va a estar condicionada por el tipo de problemas abordados en la

investigación, por el diseño utilizado para comprender y resolver dichos problemas y, entre otros aspectos, por el tipo de datos cualitativos generados. Teniendo en cuenta esta última variante, vamos a recoger aquí algunos de los métodos más conocidos y utilizados en el análisis de datos cualitativos.

Datos cualitativos textuales

Existe una gran diversidad de procedimientos que podríamos identificar como *interpretativos* (generalmente asociados a enfoques etnográficos, basados en el estudio de casos, sociológicos, etc.) para el análisis de datos cualitativos textuales, aunque la mayoría de ellos no representan sino variantes del proceso general antes descrito que se utilizan en un número reducido de investigaciones. No obstante, en algunos de estos procedimientos interpretativos se han realizado intentos de sistematización por parte de algunos autores que han contado con mayor aceptación y están ampliamente referenciadas en la literatura. Este es el caso de procedimientos como la inducción analítica, creado por Znaniecki (1934), o el método de la comparación constante de Glaser y Strauss (1967).

Uno de los métodos analíticos más característicos y sistematizados entre los investigadores interpretativos posiblemente sea la *inducción analítica*. Desde que este método fuera utilizado por Znaniecki (1934), ha sido empleado y modificado configurándose en un método riguroso, adecuado en investigaciones que se focalizan en un tema concreto. Se emplea para la generación de teoría y requiere de múltiples lugares o sujetos para su desarrollo. Consta de diversas etapas, a lo largo de las cuales se trata de ir progresivamente modificando una formulación inicial para que llegue a explicar la totalidad de los casos estudiados:

- a) Definición aproximada del fenómeno a estudiar.
- b) Formulación teórica que explique el fenómeno.
- c) Comprobación del ajuste de la hipótesis a los datos a medida que éstos van siendo recogidos.
- d) Reformulación de la hipótesis o redefinición del fenómeno en caso de no darse tal ajuste.
- e) Búsqueda deliberada de casos negativos que puedan no adaptarse a la formulación.
- f) Reformulación o redefinición en caso de hallarlos.
- g) Continuación del proceso hasta haber examinado una amplia gama de casos y haber probado adecuadamente la hipótesis.

Otro método para el análisis de datos cualitativos, también de carácter inductivo, es el desarrollado por Glaser y Strauss (1967): Estos autores, que ponen el énfasis en la generación de teoría fundamentada (basada en los datos), proponen el *método de la comparación constante* que es desarrollado mediante un procedimiento continuo en el que podemos diferenciar cuatro momentos característicos:

- a) Codificación de los datos en categorías, comparando entre sí los incluidos en

- un mismo grupo para generar propiedades teóricas de cada categoría. Simultáneamente se elaboran resúmenes recogiendo las ideas que surgen (hipótesis, relaciones, etc.).
- b) Integración de categorías y sus propiedades. La comparación de nuevos datos procedentes de los nuevos casos seleccionados (muestreo teórico) con las propiedades de la categoría, permite describir relaciones y nuevos dominios dentro de las categorías, que se van integrando en una teoría coherente.
 - c) Delimitación de la teoría y de las categorías. Cuando nuevas comparaciones producen modificaciones en la teoría, se va clarificando la lógica relacional, descartando propiedades irrelevantes, incorporando detalles, reduciendo categorías, hasta que un caso nuevo no consiga añadir nada (saturación teórica): Las teorías se reducen a conceptos de mayor nivel.
 - d) Explicitación de la teoría. Para escribirla se recurre a los resúmenes de cada categoría, se vuelve a los datos codificados para validar afirmaciones, hipótesis y proveer ilustraciones.

Basado en este enfoque se ha diseñado una aplicación informática específicamente concebida para el análisis de datos cualitativos expresados en forma de textos. Esta aplicación, conocida como NUDIST (Non-numerical Unstructured Data Indexing, Search and Theorizing), ha sido desarrollada en la Universidad de La Trobe (Australia), y será tratada con más detalle en los trabajos «NUDIST: Una herramienta informática para el análisis de datos cualitativos» (Corrales et al.), y «NUDIST versus AQUAD: Dos posibilidades informáticas para el análisis de datos cualitativos» (Rodríguez et al.) presentados ambos a este Seminario.

Otros procedimientos analíticos, no siempre suficientemente caracterizados, menos difundidos pero también utilizados con datos cualitativos de naturaleza textual, son los que proponen autores como Spradley, Goetz y Lecompte, Bogdan y Biklen, Taylor y Bogdan, Lofland y Lofland, Erickson, Miles y Huberman, y otros. Spradley (1979, 1980), interesado en describir los sistemas de significado que grupos culturales han creado y utilizan en el modo de actuar e interpretar sus experiencias, entiende el análisis como una herramienta para el descubrimiento de significado cultural a partir de los datos obtenidos mediante la entrevista en profundidad y la observación participante. El procedimiento de análisis supone la identificación de dominios (categorías), el desarrollo de taxonomías, la búsqueda de contrastes y la generación de temas (ideas síntesis), en relación con otras tantas estrategias presentes en la recogida de información.

Goetz y Lecompte (1988), quienes están interesadas por la descripción o reconstrucción analítica de los escenarios y grupos culturales a partir de las concepciones de los participantes, ponen en práctica estrategias como el análisis tipológico (clasificación en categorías), la enumeración y la aplicación de protocolos observacionales estandarizados, y procedimientos analíticos generales como los ya mencionados: método de la comparación constante e inducción analítica. Consideran, asimismo, que tanto estos procedimientos como las estrategias de selección secuencial se

apoyan en determinados modos genéricos de pensamiento y manipulación sobre los datos, a los que denominan actividades de teorización y que comprenden: a) la percepción; b) la comparación, contrastación, agregación y ordenación; c) la determinación de vínculos y relaciones; y, d) la especulación. Para estas autoras, los métodos señalados y las tareas de teorización son los instrumentos conceptuales de que dispone el investigador para desarrollar su proceso de análisis, combinándolos en cada investigación de acuerdo con sus fines y marco teórico.

Miles y Huberman (1984b) son autores de una obra dedicada específicamente al análisis de datos cualitativos. Con ella intentan aportar procedimientos sistemáticos, replicables, que permitan extraer y comprobar conclusiones. Mencionan tres grupos de técnicas constituyentes del proceso de análisis: métodos de reducción de datos (agrupamiento de códigos, sumarización, partición,...), presentación de datos (matrices, gráficos, cuadros, redes,...) y extracción y verificación de conclusiones (recuento, construcción de metáforas, detección de relaciones,... y triangulación, comprobación de la representatividad,...): Basado en el procedimiento propuesto por estos autores, Huber (1988) ha elaborado el programa AQUAD (Analysis of Qualitative Data), que responde en su concepción a los procedimientos de reducción y presentación de datos, extracción y verificación de conclusiones. Un análisis más detallado de este programa puede verse en los trabajos «Formulación y comprobación de hipótesis cualitativas con ayuda del programa AQUAD» (Gil et al.) y «NUDIST versus AQUAD: Dos posibilidades informáticas para el análisis de datos cualitativos» (Rodríguez et al.) presentados ambos a este Seminario.

Dado el coste en esfuerzo y tiempo que requiere el desarrollo de los procedimientos de análisis de datos cualitativos, se han diseñado en los últimos años un gran número de programas específicamente creados para facilitar estas tareas. Tales programas (TAP, ETHNOGRAPH, QUALPRO, QUALOG, etc.) se enmarcan en una perspectiva interpretativa y han sido ampliamente descritos en obras como las de Tesch (1988, 1989, 1990, 1991).

Datos cualitativos expresados mediante valores numéricos²

La diferenciación educativa genera datos cualitativos textuales susceptibles de transformación en valores numéricos, bien tras las operaciones de categorización y codificación, bien como resultado del recuento de unidades diferenciadas en los textos. En ambos tipos de datos (transformados mediante la categorización y codificación o procedentes del recuento de unidades en los textos) y en función del problema estudiado el investigador puede utilizar diferentes métodos de análisis.

2 Para este tipo de datos existe una serie de paquetes de programas de análisis estadístico, en los que se abordan diferentes métodos de análisis, tales como SYSTAT, SPSS, BMDP, GENSTAT, SAS... etc.

1. Datos transformados tras las operaciones de categorización y codificación

Distinguiremos en este apartado entre métodos de interdependencia entre variables y métodos de dependencia de variables. En el primer grupo de métodos de análisis, **basados en relaciones de interdependencia**, el interés se centra en la asociación mutua entre variables sin establecer distinciones (dependientes e independientes) entre ellas. Se persigue la organización y reducción de los datos en una forma más fácil de manejar por el investigador y que posibilite una mejor comprensión global de los mismos. Dentro de este grupo de métodos para el estudio de la interdependencia vamos a considerar algunas modalidades:

a) *Desarrollados a partir de tablas de contingencia*

El uso de tablas de contingencia en el análisis de datos cualitativos ha respondido a finalidades diferentes. En ocasiones, se persigue únicamente la descripción empírica de la distribución conjunta de los valores de dos variables, sin recurrir a ningún tipo de técnica estadística que permita comprender las relaciones presentes en los datos. Otras veces, no se va más allá del cálculo de alguna medida de ajuste o asociación tales como la diferencia de razones, chi-cuadrado, coeficiente de contingencia, V de Cramer, etc., medidas que sin embargo ofrecen indicaciones claras sobre la intensidad y/o el sentido de la relación entre las variables. No obstante, quizás la aplicación más atractiva de este tipo de tablas sea la que las vincule al desarrollado procedimientos analíticos de mayor complejidad, para datos de naturaleza secuencial y no secuencial. Este es el caso del análisis secuencial, el sistema de diferencia de proporciones, el análisis lineal-logarítmico o el análisis factorial confirmatorio.

Las *técnicas secuenciales* se utilizan para solucionar problemas de ordenación de los datos en el tiempo, lo que a su vez permite comprender cómo esos datos se producen momento a momento. En este sentido, el análisis secuencial recoge datos relativamente momentáneos y no centrados en un desarrollo que se extiende a lo largo de meses o años: los datos secuenciales recogen la continuidad entre unidades sucesivas codificadas (Bakeman y Gottman, 1989:80): El análisis secuencial parte de la obtención de una serie de estadísticos descriptivos simples (duración media y probabilidad simple del evento, probabilidad de transición), calculados a partir de una tabla de contingencia en la que cada celda indica el número de veces que ocurre un determinado evento dentro de una secuencia. Estos valores se comparan con modelos teóricos, equiprobables o no, a partir de estadísticos como chi-cuadrado de bondad de ajuste o la prueba binomial de puntuaciones z. Aunque el procedimiento es relativamente simple y puede ser desarrollado de forma manual, existen programas informáticos como ELAG que nos facilitan el desarrollo de este tipo de análisis (Bakeman, 1983).

El *sistema de diferencia de proporciones* se aplica al estudio de tablas de contingencia multidimensionales y se desarrolla de acuerdo con tres ideas básicas:

la medida base de asociación es la diferencia de proporciones o porcentajes; parte del «paradigma de la elaboración», desarrollado por Lazarsfeld (1955) y Rosenberg (1968), asignando además valores a las relaciones conceptuales del paradigma; sigue los principios del gráfico de flujo, para representar los sistemas de ecuaciones lineales en forma de grafos. El análisis, siguiendo el sistema de diferencia de proporciones, o sistema de la D, descompone una tabla de contingencia multidimensional en una serie de ecuaciones para más tarde representar éstas en forma de grafos. El procedimiento que se sigue para estimar los coeficientes de relación «d» en el grafo supone calcular proporciones y estandarizarlas. Estos coeficientes «d» son susceptibles de ser sometidos a prueba estadística calculando su intervalo de confianza (Sánchez Carrión, 1984): Un programa informático que permite el análisis de la diferencia de proporciones es el programa CHIP (Bogart y Conner, 1986).

Las técnicas denominadas *modelos lineales logarítmicos* (Kennedy, 1983) también permiten al investigador estudiar tablas de contingencia multidimensionales, midiendo y contrastando las asociaciones e interacciones que aparecen en ellas. Se basan en un modelo aditivo, de acuerdo con el cual es posible determinar el logaritmo en base e de la frecuencia esperada en cada casilla a partir de los efectos principales y debidos a las interacciones entre las variables que forman la tabla. Constituyen una respuesta satisfactoria al análisis de los datos relativos a variables categóricas obtenidas tras los procesos de reducción de datos cualitativos en forma de registros narrativos, citas textuales de sujetos o documentos (Anguera, 1985).

También es posible desarrollar procedimientos de análisis factorial booleano y confirmatorio. En el primer caso, el *análisis factorial booleano* es una alternativa al análisis factorial clásico para variables dicotómicas basadas en el álgebra de Boole. Su objetivo es expresar un conjunto de variables observadas por medio de un grupo menor de factores o variables latentes. La obtención de los factores se realiza a partir del cálculo de matrices de discrepancia entre el valor observado y el estimado. Las puntuaciones factoriales de un individuo tienen valor uno si éste presenta una discrepancia para cualquiera de las variables dominantes en el factor (las que no tienen una saturación cero) y cero en el caso contrario.

Por su parte, el *análisis factorial confirmatorio* es una técnica que nos permite descubrir una serie de variables latentes estudiando la covariación entre un grupo de variables observadas, y atendiendo a una serie de restricciones: los factores comunes no tienen por qué estar necesariamente correlacionados; las variables observadas no precisan tener un factor único asociado a ellas; es posible encontrar correlaciones entre los factores únicos. Cuando se aplica a datos cualitativos, el análisis se inicia con la construcción de una tabla de contingencia a partir de la cual se calcula la correlación tetracórica entre las variables. Este tipo de correlación implica realizar una estimación de la correlación entre variables latentes que presentan una distribución conjunta normal bivalente. La estimación puede desarrollarse por diferentes métodos. Los programas PRELIS y LISREL7 (Jöreskog y Sörbom, 1989) facilitan los cálculos necesarios para la contrastación del modelo diseñado en cada caso.

b) *Orientados a la descripción mediante relaciones topológicas*

La dificultad de asociar valores numéricos que representen isomórficamente a las realidades educativas ha llevado a considerar métodos estadísticos basados en medidas ordinales o nominales, y generalmente interesados por modelos de representación gráficos más que por la expresión algebraica numérica. Como alternativa, se ha desarrollado una línea de análisis estadísticos que se apoya en las relaciones ordinales entre los objetos; lo esencial en el análisis no es la precisión numérica sino la estructura de un sistema de posiciones y distancias relativas entre objetos que es reflejado mediante un lenguaje geométrico-espacial (Conde, 1987): Entre ellos, pueden citarse los métodos factoriales desarrollados a partir de las aportaciones de Benzécri (1973) y los métodos de escalamiento no métrico.

Los métodos factoriales operan a través de cálculos basados en el álgebra lineal, y proporcionan representaciones gráficas en las que se sintetiza un amplio número de valores numéricos, situando los objetos a describir sobre ejes factoriales. El antecedente de estas técnicas se encuentra en el análisis factorial clásico, pero a diferencia de éste, en el que se considera que cada sujeto puede ser resultado de la combinación de unos factores comunes, unos factores específicos y una perturbación aleatoria, en los procedimientos de la escuela francesa no se parte de un modelo a priori. Entre las técnicas más empleadas se encuentra el *análisis de correspondencias*, el cual proporciona al analista una representación geométrica de la información contenida en una matriz de datos que adopta la forma de una tabla de contingencia. Los resultados de un análisis de correspondencias se presentan en gráficos que indican las configuraciones de puntos en proyecciones planas formadas por los ejes principales tomados de dos en dos (Lebart, Morineau y Warwick, 1984:44): La inspección de estos gráficos permitiría identificar los rasgos que se corresponden con cada eje y los puntos-fila y/o puntos-columna agrupados en una misma región del espacio. Existen un conjunto de paquetes de programas que facilitan este método de análisis desde diferentes perspectivas, tales como el SPAD-N (Lebart, Morineau y Lambert, 1988), ANCORSIM (Cornejo, 1988), la serie Dual3 (Nishisato y Nishisato, 1984), Mapwise (Goodnow, 1988), los programas CORAN y CORRESP del paquete PC-MDS (Carroll et al., 1986) o el programa SimCA (Greenacre, 1986).

El *escalamiento multidimensional* incluye una variedad de técnicas cuyo propósito común es hacer explícito el modelo o estructura que se encuentra oculto en una matriz de datos, y representar esa estructura en algún modo que resulte mucho más accesible para el ojo humano (Shepard, Rommey y Nerlove, 1972): Estas técnicas operan sobre una matriz cuadrada de distancias entre objetos o individuos para obtener una representación espacial consistente en una configuración geométrica de puntos, a modo de mapa, que refleja la estructura oculta de los datos y hace que las relaciones entre éstos sean más fáciles de comprender (Kruskal y Wish, 1986): En el escalamiento multidimensional *no métrico*, las distancias calculadas entre cada par de objetos son números que expresan una similitud o diferencia entre ambos. El

objetivo de este procedimiento es encontrar un conjunto de puntos en un espacio de dimensión mínima tales que los datos de disimilaridad sean una función monótona (ordenada) de las distancias en ese espacio. Es decir, para todo conjunto de datos que estén ordenados del mismo modo siempre obtendremos idéntica solución (Coxon y Jones, 1984): El programa MDSCAL de Kruskal fue el primero de los creados para llevar a cabo el escalamiento multidimensional no métrico. Entre los programas más recientes se encuentra, con el mismo nombre, el elaborado por Orloci y Kenkell (1987).

c) *Basados en procedimientos de clasificación*

Se trata de técnicas que persiguen la asignación de un nuevo objeto u observación a su propio lugar en un conjunto bien establecido de categorías, conglomerados o clases. Cuando trabajamos con datos cualitativos, podemos realizar clasificaciones de naturaleza estadística mediante el empleo de métodos multivariantes tales como los análisis de conglomerados o las técnicas de clasificación de la escuela francesa. Los primeros, los métodos denominados *análisis de conglomerados*, tienen por finalidad agrupar objetos o individuos que son similares o tienen algunas características comunes. Se parte de una matriz de n sujetos sobre los que se han medido p variables. Esta matriz se transforma en otra de dimensiones $n \times n$ donde se recogen las semejanzas (o desemejanzas) entre cada pareja de objetos, de acuerdo con un criterio de proximidad (o lejanía): A partir de ella se forman los grupos. Dunn-Rankin (1983:122) distingue cuatro pasos en el proceso del análisis de conglomerados: recogida de la información sobre los objetos, obtención de las distancias entre todos los objetos, agrupamiento de los objetos en base a una medida de proximidad (o lejanía) y representación gráfica de los grupos que han sido formados.

Los *métodos de clasificación* desarrollados por la escuela francesa de análisis de datos son similares a los análisis de conglomerados mencionados más arriba. Permiten agrupar en clases o jerarquías de clases objetos o individuos sobre los cuales poseemos cierta información. Utilizan procedimientos basados en cálculos algorítmicos y producen representaciones gráficas generalmente en forma de árbol invertido. Lebart, Morineau y Fenelon (1985) presentan diversas técnicas de clasificación que emplean *algoritmos ascendentes o aglomerativos*, en los cuales se procede a la construcción de las clases por aglomeraciones sucesivas de elementos dos a dos, dando lugar a una jerarquía de la partición de los elementos; *algoritmos descendentes o divisivos*, que proceden mediante dicotomías sucesivas del conjunto de los elementos, propiciando igualmente una jerarquía de la partición; y algoritmos que conducen directamente a las *particiones*. Como programa informático para el desarrollo de los procedimientos de clasificación podemos utilizar SPAD-N (Lebart, Morineau y Lambert, 1988):

Las técnicas **basadas en relaciones de dependencia** consideran previamente variables explicativas y variables a explicar. Por tanto, resultan adecuadas cuando

interesa predecir determinadas variables en función de otras. Básicamente, los dos métodos para el estudio de la dependencia entre variables utilizados con datos cualitativos son los modelos logit y el análisis de correlación canónica. Los *modelos logit* se interesan por el estudio del efecto de un conjunto de factores (variables independientes) sobre una variable dependiente considerada como respuesta. En este análisis está presente la idea de la regresión múltiple. A diferencia de los modelos lineales logarítmicos, cuando se ajusta un modelo logit se consideran todas las asociaciones e interacciones significativas que incluyen la variable dependiente, junto con la interacción entre las variables independientes. Se trata de encontrar qué efectos de los que incluyen la variable dependiente son significativos (Sánchez Carrión, 1984:287-8).

El *análisis de correlación canónica*, por su parte, es adecuado cuando pretendemos predecir los valores de un conjunto de variables (criterios) en función de los obtenidos para otro conjunto de ellas (predictores): Se trata de obtener dos combinaciones lineales, una de las variables criterio y otra de las variables predictores, tales que la correlación simple entre la variable independiente compuesta y la variable dependiente compuesta sea máxima (Warwick, 1975): El uso del análisis de correlación canónica con propósitos descriptivos no requiere asunciones sobre la distribución de las variables; en tales casos, las variables predictoras y criterios pueden ser medidas en el nivel nominal u ordinal (Dillon y Goldstein, 1984:339).

2. Datos transformados procedentes del recuento de unidades diferenciadas en los textos

Como hemos expresado anteriormente, una de las vías habituales en el análisis de datos cualitativos textuales consiste en la aplicación de técnicas estadísticas a los valores numéricos resultantes del recuento de unidades en los textos. El análisis interpretativo de los datos suele verse complementado con un recuento de las unidades incluidas en cada categoría considerada durante el proceso analítico. A veces, la cuantificación se utiliza únicamente para valorar la importancia de determinados tópicos o conceptos dentro del corpus de datos, pero es posible también construir tablas de contingencia cruzando las categorías con distintos sujetos, instituciones, contextos, etc. estudiados. En este caso, los datos resultantes de la cuantificación de los textos pueden ser objeto de las técnicas que se han presentado ya al hablar de los datos cualitativos transformados en valores numéricos tras las operaciones de categorización y codificación.

Sin embargo, en los últimos años han cobrado un impulso considerable los enfoques analíticos que se basan en la cuantificación directa de los textos, sin que medie una reducción o transformación previa de los datos por parte del analista. El procedimiento consiste en contar las unidades elementales presentes en los textos (palabras o secuencias breves de palabras) y realizar análisis estadísticos a partir de los recuentos. El análisis estadístico de textos tiene su origen en los análisis cuantitativos realizados sobre obras literarias, que iban dirigidos al recuento de palabras,

el estudio de la distribución del vocabulario, la comparación del léxico empleado por distintos autores o por un mismo autor en diferentes períodos creativos. Las investigaciones de autores como Yule (1944), Zipf (1946), Guiraud (1960), Muller (1968), entre otros, y el posterior desarrollo y popularización de la informática se encuentran en la base de los métodos de la denominada estadística textual, que han acabado aplicándose al estudio de los datos textuales en muy diversos ámbitos: historia, literatura, sociología, etc.

La aparición del programa GENERAL INQUIRER (Stone y otros, 1966) marcó el inicio del análisis automático de textos. El programa se basaba en la búsqueda y recuento de palabras y frases previamente identificadas por el analista mediante la definición de un diccionario confeccionado para el análisis. Mochmann (1983) describe programas más recientes basados en estos mismos principios: EVA, SPENCE, COFTA, COTAG, TEXPACK. Otros programas, en lugar de partir de un diccionario previo, han tratado de extraer los temas presentes en un texto sometiendo a tratamientos estadísticos las frecuencias de cada palabra en cada una de las unidades consideradas. Es el caso del programa WORDS (Iker, 1975): La idea base sigue siendo el despiece automático de un texto en unidades elementales, para operar cálculos estadísticos a partir de su recuento.

Los métodos de la escuela francesa de análisis de datos, resultan especialmente adecuados para el análisis de grandes matrices de datos como las originadas al examinar la distribución de unidades elementales dentro de un texto. Enmarcados en esta línea, se han difundido programas específicamente diseñados para el análisis de datos textuales, tales como LEXICLOUD (Salem, 1987), ALCESTE (Reinert, 1986) o SPAD.T (Lebart, Morineau y Bécue, 1989): Las técnicas que permiten desarrollar se basan, en todos ellos, en el análisis de correspondencias, las clasificaciones jerárquicas o el cálculo de medidas de asociación (chi-cuadrado, método de las especificidades):

El procedimiento general de análisis de datos cualitativos así como los métodos específicos en que éste puede concretarse, representan propuestas útiles para reducir, disponer y obtener conclusiones significativas a partir de la evidencia recogida en relación con los diferentes problemas a los que se enfrenta la investigación sobre la diferenciación educativa. Las técnicas presentadas para el análisis de datos numéricos, procedentes de la transformación de datos textuales, han sido empleadas para abordar algunas de las cuestiones que, desde hace años, vienen preocupando a los investigadores que trabajan en este campo. Partiendo de datos cualitativos numéricos, para la identificación de rasgos se recomienda el análisis factorial o para la construcción de tipologías el análisis de conglomerados (Pérez Juste, 1985): Por otra parte, a un nivel más general, Tejedor (1988a) plantea la utilización del análisis factorial de correspondencias o del escalamiento multidimensional para describir o reducir datos; el uso del análisis canónico con una intención predictiva; o la utilidad del análisis de conglomerados para agrupar la información recogida.

No obstante, queremos insistir también en el interés que presentan las técnicas

de análisis de datos textuales, no basadas en procedimientos estadísticos. Entendemos que las principales aportaciones del análisis de datos cualitativos al campo de la diferenciación educativa cabe situarlas en la resolución de problemas en los que el pluralismo cultural y lingüístico intervengan como elemento fundamental. Del mismo modo, en aquellos problemas en los que las categorías culturales o significados propios de cada grupo social resulten variables fundamentales para explicar o intervenir en determinadas situaciones educativas. En estos casos, resulta más apropiado un tipo de análisis que mantenga en esencia la estructura y los términos que utiliza un individuo o una comunidad para expresar y comunicar su enfoque de la realidad.

El análisis de datos cualitativos puede permitir también aproximarse con éxito a aquellas investigaciones o intervenciones educativas preocupadas por reflejar las interacciones propias del proceso de enseñanza-aprendizaje para los diferentes niveles educativos o en función de las diferencias debidas al sexo. Este tipo de análisis resulta adecuado para estudiar modelos de interacción social y lingüística, como los que se desarrollan entre alumnos y profesores en las clases o en situaciones educativas informales.

Este tipo de análisis nos permite asimismo profundizar en las características idiosincráticas de cada centro educativo en su lucha por adaptar la educación a las características contextuales donde se inserta. Recoger la cultura y el clima de relaciones propios del centro, explicar las múltiples interacciones que tienen lugar dentro de las instituciones educativas, y las que éstas mantienen con los diferentes elementos del contexto en que se insertan, valorar los cambios curriculares introducidos en un centro y/ o en una clase, son algunos de los problemas para los que el análisis de datos cualitativos puede ofrecer respuestas satisfactorias. De igual modo, este tipo de análisis nos ayuda a profundizar en los matices, contradicciones, anécdotas, peculiaridades o particularidades del discurso de personas, grupos o instituciones implicadas en acciones educativas, y que constituyen puntos de referencia fundamentales para explicar ciertas diferencias.

Una aplicación más del análisis de datos cualitativos está vinculada con la explicación global de las diferencias educativas. Este tipo de análisis nos permite alcanzar una comprensión ecológica de los fenómenos, de modo que sea posible entender las relaciones entre un sujeto (grupo, institución, etc.) y su medio, y las relaciones entre los diferentes medios en los que vive ese sujeto. Elementos tales como su historia personal (y la del propio medio en que vive), su enfoque de la realidad (y el mantenido por el grupo social del que es parte), su conductas escolares y no escolares, etc. pueden interpretarse holísticamente bajo formas de análisis de datos cualitativos.

REFERENCIAS BIBLIOGRÁFICAS

- ALBERDI, I. (Coord.) (1985): *Mujer y educación*. Madrid, Instituto de la Mujer. Serie Documentos 3.
 ALBERDI, I. (1986): La educación de la mujer en España, en C. BORREGUERO: *La mujer española: de la tradición a la modernidad*, Madrid, Tecnos.

- ALBERDI, I. y ESCARIO, P. (1987): Sexismo en la enseñanza media, en INSTITUTO DE LA MUJER. *La investigación en España sobre mujer y educación*. Madrid, Instituto de la Mujer.
- ANDER-EGG, E. (1990): *Técnicas de Investigación Social*. Buenos Aires, Editorial Humanitas.
- ANGUERA, M. T. (1985): Posibilidades de la metodología cualitativa vs. cuantitativa. *Revista de Investigación Educativa*, 3 (6), 127-144.
- BAKEMAN, R. (1983): Computing Lag Sequential Statistics: The ELAG Program. *Behavior Research Methods and Instrumentation*, 15, 530-535.
- BAKEMAN, R. y GOTTMAN, J. M. (1989): *Observación de la interacción: Introducción al análisis secuencial*. Madrid, Morata.
- BANKS, J. (1993): Multicultural education: Historical development, dimensions and practice, *Review of Research in Education*, 19, 3-49.
- BARTOLOMÉ PINA, M. (1992): Diseños y metodologías de investigación desde la perspectiva de la educación intercultural. En X Congreso Nacional de Pedagogía. *Educación Intercultural en la Perspectiva de la Europa Unida*. Salamanca, (pp. 647-674).
- BENZECRI, J. P. (1973): *L'Analyse des Données*. París, Dunod.
- BLAIKIE, N. W. (1991): A critique of the use of triangulation in social research. *Quality & Quantity*, 25, 115-136.
- BLISS, J., MONK, M. y OGBORN, J. (1983): *Qualitative Data Analysis for Educational Research*. Londres, Croom Helm.
- BOGDAN, R. C. (1986): A sociology of special education, en R. J. MORRIS y B. BLATT (Eds.): *Special education: Research and trends*. Oxford, Pergamon, (pp. 334-359).
- BOGDAN, R. C. y BIKLEN, S. K. (1982): *Qualitative Research for Education. An Introduction to Theory and Methods*. Boston, Allyn and Bacon, Inc.
- BOWMAN, B. (1993): Early Childhood Education, *Review of Research in Education*, 19, 101-134.
- BRULLET, C. y SUBIRATS, M. (1990): *La coeducación*. Madrid, MEC
- CARTWRIGHT, D. P. (1978): Análisis del Material Cualitativo, en L. FESTINGER y D. KATZ (Eds.): *Los Métodos de Investigación en las Ciencias Sociales*. Buenos Aires, Paidós.
- CARROLL, J. D. et al. (1986): Interpoint Distance Comparisons in Correspondence Analysis. *Journal of Marketing Research*, (23), 271-280.
- CASELLES, J. F. (1991): Pedagogía diferencial, sexo y educación: El sexismo en el sistema educativo mixto y la coeducación como alternativa. En C. JIMÉNEZ FERNÁNDEZ, *Lecturas de Pedagogía Diferencial*. Madrid: Dykinson, (pp. 105-134).
- CORRALES, A. et al.: «NUDIST»: Una herramienta informática para el análisis de datos cualitativos. Comunicación presentada al VI Seminario de Métodos de Investigación, Madrid.
- CONDE, F. (1987): Una propuesta de uso conjunto de las técnicas cuantitativas y cualitativas en la investigación social. El isomorfismo de las dimensiones topológicas de ambas técnicas. *Revista Española de Investigaciones Sociológicas*, (39), 213-224.
- COXON, A. P. y JONES, C. L. (1984): Escalas multidimensionales, en J. J. SÁNCHEZ CARRIÓN *Introducción a las técnicas de análisis multivariable aplicadas a las ciencias sociales*. Madrid, Centro de Investigaciones Científicas, pp. 107-129.
- DAWSON, J. A. (1982): *Qualitative research findings: what do we do to improve and estimate their validity?* Comunicación presentada al Annual Meeting of the American Educational Research Association, Nueva York. ERIC Document Reproduction Service N° ED218330.
- DENZIN, N. K. (1978): *The research act*. Nueva York, McGraw Hill.
- DILLON, W. R. y GOLDSTEIN, M. (1984): *Multivariate analysis*. Nueva York, John Wiley & Sons.
- DUNN-RANKIN, P. (1985): *Scaling methods*. Hillsdale, New Jersey, Lawrence Erlbaum Associates, Publishers.

- ERICKSON, F. (1989): Métodos Cualitativos de Investigación sobre la Enseñanza, en M. C. WITTRICK (Ed.): *La Investigación de la Enseñanza, II*. Madrid, Paidós-MEC, pp. 125-301.
- ERICKSON, F. (1982): Audiovisual Records as a Primary Data Source, en A. GRIMSHAW (Ed.): *Sociological Methods and Research (Special Issue on Sound-Image Records in Social Interaction Research)*, 11, 2, 213-232.
- ESPÍN, J. V. (1991): Los programas de educación compensatoria. ¿Una respuesta a las diferencias socioculturales desde la educación?, en C. JIMÉNEZ FERNÁNDEZ: *Lecturas de Pedagogía Diferencial*. Madrid, Dykinson, pp. 105-134.
- FIELDING, N. G. y FIELDING, J. L. (1986): *Linking data*. Sage University Paper serie on Qualitative Research Methods, 4. Beverly Hills, Sage Publications.
- FILLMORE, L. W. y VALADEZ, C. (1986): Teaching Bilingual Learners. En WITTRICK, M. C. (Ed.): *Handbook of Research on Teaching*. Nueva York: MacMillan.
- FIRESTONE, W. A. y DAWSON, J. A. (1982): *Approaches to qualitative data analysis: intuitive, procedural and intersubjective*. Comunicación presentada al Annual Meeting of the American Educational Research Association. Nueva York.
- FRANZOSI, R. (1990): Strategies for the prevention, detection, and correction of measurement error in data collected from textual sources. *Sociological Methods and Research*, 18 (4), 442-472.
- GAVIRIA SOTO, J. L. (1986): El enfoque diferencial en la investigación pedagógica. *Bordón*, (262), 199-221.
- GIL, J. et al.: *Formulación y comprobación de hipótesis cualitativas con ayuda del programa AQUAD*. Comunicación presentada al VI Seminario de Métodos de Investigación, Madrid.
- GLASER, B. G. y STRAUSS, A. L. (1967): *The Discovery of Grounded Theory*. Nueva York, Aldine Publishing Company.
- GOETZ, J. P. y LECOMPTE, M. D. (1988): *Etnografía y Diseño Cualitativo en Investigación Educativa*. Madrid, Morata.
- GOODNOW, W. E. (1988): *Brief User's Manual*. Chicago, IL: Research Software Inc.
- GREENACRE, M. J. (1986): SimCA: A Program to Perform Simple Correspondence Analysis. *The American Statistician*, (51), 230-231.
- GUBA, E. G. (1983): Criterios de credibilidad en la investigación naturalista, en A. PÉREZ y J. GIMENO, (Eds.): *La enseñanza: su teoría y su práctica*. Madrid, Akal, pp. 148-165.
- GUIRAUD, P. (1960): *Problèmes et méthodes de la statistique linguistique*. París, PUF.
- HAMMERSLEY, M. y ATKINSON, P. (1983): *Ethnography: principles and practice*. Londres, Tavistock.
- HEGARTY, S. (1988): Qualitative research: Introduction, en S. HEGARTY y P. EVANS (Eds.): *Research methods in special education*, Oxford, NFER-Nelson, (pp. 109-113).
- IKER, H. P. (1975): *Words, system manual*. Rochester, Nueva York, Computer Print.
- IKER, H. P. y HARWAY, N. J. (1969): A computer systems approach toward the recognition and analysis of content, en G. A. GERBNER y otros (Eds.): *The analysis of communication content*. Nueva York, 381-405.
- JACOB, E. (1988): Clarifying qualitative research: a focus on traditions. *Educational Researcher*, 17 (1), 16-24.
- JÖRESKOG, K. G. y SÖRBOM, D. (1989): *Lisrel 7. User's Reference Guide*. Mooresville, IN, Scientific Software, Inc.
- KENNEDY, J. J. (1983): *Analyzing Qualitative Data. Introductory Log-Linear Analysis for Behavioral Research*. Praeger.
- KERLINGER, F. N. (1975): *Investigación del Comportamiento*. México, Nueva Editorial Interamericana.
- KRUSKAL, J. B. y WISH, M. (1986): *Multidimensional scaling*. Sage University Paper Series on Quantitative Applications in the Social Sciences, 11. Beverly Hills, Sage Publications.

- LAZARSELD, P. F. y ROSENBERG, M. (1955): *The language of social research: A reader in the methodology of social research*. Nueva York: Free Press.
- LEBART, L., MORINEAU, A. y FENELON, J. P. (1985): *Tratamiento Estadístico de Datos*. Barcelona, Marcombo.
- LEBART, L., MORINEAU, A. y LAMBERT, T. (1988): *SPAD.N: Système Portable pour l'Analyse des Données*. París: CISIA
- LEBART, L., MORINEAU, A. y WARWICK, K. M. (1984): *Multivariate Descriptive Statistical Analysis: Correspondence Analysis and Related Techniques for Large Matrices*. Nueva York: John Wiley.
- LEVINE, H. G. (1985): Principles of data storage and retrieval for use in qualitative evaluations. *Educational Evaluation and Policy Analysis*, 7 (2), 169-186.
- LOFLAND, J. y LOFLAND, L. H. (1984): *Analyzing social settings*. Belmont, Wodsworth, Publishing Company.
- MCINTYRE, D. y MACLEON, G. (1986): The characteristics and uses of sistematic classroom observation, en M. HAMMERSLEY (Ed.): *Controversies in classroom research*, Milton Keynes, Open University Press, 10-24.
- MILES, M. B. (1985): Qualitative data as an attractive nuisance: the problem of analysis, en J. VAN MAANEN (Ed.): *Qualitative Methodology*. Beverly Hills, Sage Publications, pp. 117-134.
- MILES, M. B. y HUBERMAN, A. M. (1984a): Drawing valid meaning from qualitative data: toward a shared craft. *Educational Researcher*, 13 (4), 20-30.
- MILES, M. B. y HUBERMAN, A. M. (1984b): *Qualitative Data Analysis. A Sourcebook of New Methods*. Beverly Hills, Sage Publications.
- MOCHMANN, E. (1985): Análisis de Contenido mediante Ordenador Aplicado a las Ciencias Sociales. *Revista Internacional de Sociología*, 43 (1), 11-44.
- MOHLER, P. Ph. (1985): Algunas Observaciones Prácticas sobre la Utilización del Análisis de Contenido en Ordenador. *Revista Internacional de Sociología*, 43 (1), 45-57.
- MOLLY, E. (1984): *Factors affecting the reduction of narrative data*. Comunicación presentada al Annual Meeting of the American Educational Research Association, New Orleans. ERIC Document Reproduction Service N° ED247300.
- MUCCHIELLI, R. (1988): *L'analyse de contenu des documents et des communications*. París, Les Editions ESF-Entreprise Moderne d'Édition.
- MULLER, C. (1968): *Initiation à la statistique linguistique*. París, Larousse.
- MULLER, C. (1977): *Principes et methodes de statistique lexicale*. París, Hachette.
- ORLOCI, L. y KENKEL, N. C. (1987): *Data analysis in population and community ecology*. Draft Copy, Universidad de Hawaii, Honolulu & New Mexico State University, Las Cruces.
- PÉREZ JUSTE, R. (1985): *Elementos de Pedagogía Diferencial*, Madrid, UNED.
- REINERT, M. (1986): Un logiciel d'analyse lexical (ALCESTE): *Les Cahiers de l'Analyse des Données*, XI (4), 471-484.
- RODRÍGUEZ, G. et al.: *NUDIST versus AQUAD: Dos posibilidades informáticas para el análisis de datos cualitativos*. Comunicación presentada al VI Seminario de Métodos de Investigación, Madrid.
- RODRÍGUEZ, G. (1992): La investigación sobre la eficacia escolar. En C. GARCÍA PASTOR: *La investigación sobre la integración: Tópicos, aproximaciones y procedimientos*. Salamanca, Amarú.
- ROSENBERG, M. (1968): *The logic of survey analysis*. Nueva York: Basic Books.
- SALEM, A. (1987): *Le Lexicloud. Programmes pour le traitement lexicométrique des textes*. Ecole Normale de Fontenay-Saint Cloud.
- SÁNCHEZ CARRIÓN, J. J. (1984): *Introducción a las técnicas de análisis multivariados aplicados a las ciencias sociales*. Madrid, Centro de Investigaciones Sociológicas.

- SÁNCHEZ CARRIÓN, J. J. (1985): Técnicas de Análisis de los Textos mediante Codificación Manual. *Revista Internacional de Sociología*, 43 (1), 89-118.
- SÁNCHEZ CARRIÓN, J. J. (1989): *Análisis de tablas de contingencia. El uso de los porcentajes en las ciencias sociales*. Madrid, Centro de Investigaciones Sociológicas - Siglo XXI de España Editores, S.A.
- SHEPARD, R. N.; ROMNEY, A. K. y NERLOVE, J. B. (1972): *Multidimensional scaling. Theory and applications in the behavioral sciences*. Nueva York, Seminar Press.
- SMITH, J. K. y HESHUSIUS, L. (1986): Closing down the conversation: the end of the quantitative-qualitative debate among educational inquirers. *Educational Researcher*, 15 (1), 4-12.
- SPRADLEY, J. P. (1979): *The ethnographic interview*. Nueva York, Holt, Rinehart and Winston.
- SPRADLEY, J. P. (1980): *Participant observation*. Nueva York, Holt, Rinehart and Winston.
- STONE, P. J. y otros, (1966): *The General Inquirer: a computer approach to content analysis*. Cambridge, Massachusetts, the M.I.T. Press.
- STRAUSS, A. L. (1987): *Qualitative analysis for social scientists*. MA, Cambridge University Press.
- STRAUSS, A. y CORBIN, J. (1990): *Basics of qualitative research. Grounded theory procedures and techniques*. Newbury Park, Sage Publications.
- TEJEDOR, F. J. (1988a): La Pedagogía Diferencial en su conceptualización metodológica. *Bordón*, 40 (4), 589-598.
- TEJEDOR, F. J. (1988b): *Perspectivas de la investigación cuantitativa en educación*. Cádiz. Documento policopiado.
- TESCH, R. (1988): *The impact of the computer on qualitative data analysis*. Paper presented at the Annual Meeting of the American Educational Research Association.
- TESCH, R. (1989): Basic Qualitative Analysis with QUALPRO. *Qualitative Studies in Education*, 2 (4), 367-371.
- TESCH, R. (1990): *Qualitative Research: Analysis Types and Software Tools*. Bristol, The Falmer Press.
- TESCH, R. (1991): Software for qualitative researchers: analysis needs and program capabilities, en N. G. FIELDING y R. M. LEE (Eds.): *Using computers in qualitative research*. Londres, Sage Publications, pp. 16-37.
- VAN MAANEN, J. (1985): The fact of fiction in organizational ethnography, en J. VAN MAANEN (Ed.): *Qualitative Methodology*, Beverly Hills, Sage Publications, pp. 37-55.
- WAINER, H. (1992): Understanding graphs and tables. *Educational Researcher*, 21 (1), 14-23.
- WARWICK, P. V. (1975): Canonical correlation analysis, en N. H. NIE et al.: *SPSS: Statistical package for the social sciences*, Nueva York, McGraw Hill, pp. 515-527.
- WEBER, G. (1971): *Inner-city children can be taught to read: Four succesful schools*. Washington DC: Council for Basic Education.
- WILSON, S. (1977): The use of ethnographic techniques in educational research. *Review of Educational Research*, 47 (1), 245-265.
- WOLCOTT, H. F. (1986): *Problem finding in qualitative research*. Universidad de Oregon, Documento de trabajo.
- WOODS, P. (1989): *La escuela por dentro*. Barcelona, Paidós-MEC.
- YULE, G. U. (1944): *A statistical study of vocabulary*. Cambridge, University Press.
- ZIPF, G. K. (1946): *The psychobiology of language, an introduction to dynamic philology*. Boston, Houghton-Mifflin.
- ZNANIECKI, F. (1934): *The method of sociology*. Nueva York, Farrar & Rinehart.

COMUNICACIONES

COMUNICACIONES PONENCIA I

EVALUANDO LA EDUCACIÓN INFANTIL: REVISIÓN DEL MODELO Y METODOLOGÍA

Mercedes García García

Esta comunicación es el resultado de la primera fase de un proyecto sobre el estudio de la eficacia diferencial de la Educación Infantil (EI). Nuestra finalidad principal es elaborar un modelo que, teniendo en cuenta las diferencias individuales, optimice el desarrollo y el aprendizaje del niño preescolar así como nos permita predecir el tipo de intervención educativa más adecuada para cada sujeto. Como síntesis de los resultados de la investigación sobre efectos y eficacia de la EI y de una concepción adaptativa de la educación elaboramos una primera aproximación a un modelo educativo.

La elaboración y validación empírica de este tipo de modelos encuentra graves problemas debido a las limitaciones teóricas y empíricas halladas en la evaluación de la EI. Las limitaciones más importantes se deben a deficiencias metodológicas, en concreto podemos hacer referencia a: 1) la relativa naturaleza atórica de la investigación sobre EI, dando como resultado pequeñas parcelas independientes de saber, cierto grado de inconsistencia entre problema y método de estudio, puesto que no existe un marco coherente donde integrar los resultados de los estudios realizados o validar los supuestos teóricos subyacentes; 2) el tipo de población utilizada puesto que la mayor parte de los estudios se refieren a estudiantes con algún tipo de déficit al ser, en su gran mayoría, resultados derivados de las evaluaciones del proyecto Head Start, siendo escasos los estudios con poblaciones regulares y, en la mayoría de las ocasiones, poco representativos; 3) el enfoque de estudio: se han realizado críticas a la debilidad del diseño utilizado debido, principalmente, a un intento de seguir las reglas de experimentalidad cuando la EI —como cualquier otro tipo de intervención educativa— necesita ser analizada en la complejidad del aula, en un enfoque multivariado y de cambio para ser comprendida en su globalidad; y 4) todas aquellas limitaciones referidas a la falta de descripción operativa de los niveles de la/s variable/s independiente/s; en torno a la medida de la/s variable/s dependiente/s; o la escasa adecuación de los análisis estadísticos utilizados, etc.

El primer modelo del que partíamos constaba de cuatro grandes elementos relacionados: 1) *Población*, referido a aquellas variables relativas a diferencias de entrada (input), grupales e individuales, que parecen relacionarse a los diferentes resultados (output) de los niños que han recibido intervención preescolar; 2) *Programa*, referido a aquellas relativas a diferencias en la intervención preescolar y que también parecen relacionarse a los resultados; 3) *Padres*, referido a un elemento que, aunque externo a la escuela parece tener una especial relevancia en este nivel educativo por la mayor implicación y participación en el programa preescolar; y 4) *Competencia individual*, referido a las tradicionales variables de resultado o características infantiles relativas al desarrollo (cognitivo y socioafectivo) y a las variables procesuales de aprendizaje del niño preescolar que se intentan modificar después o

durante una intervención educativa y que podrían estar, a su vez, relacionados a los objetivos de los programas educativos. Al mismo tiempo, los resultados logrados pudieran influir en cada uno de los otros elementos modificándolos.

A partir de este modelo se empezó a trabajar en la **Primera Fase** de nuestro proyecto con un doble objetivo: 1) Chequear teóricamente la consistencia de los elementos generales del modelo, su operativización y sus relaciones y 2) Encontrar una metodología de investigación que tuviera en cuenta el proceso de enseñanza-aprendizaje en contextos reales dentro de un enfoque multivariado y de cambio individual.

Objetivo 1º: Chequear los elementos del modelo

1) **POBLACIÓN**: Parece que no es conveniente hablar de conductas de intervención eficaz sin especificar la **edad** del niño. Si se considerara, pues, como grupo parece haber dos grupos: de 10 meses a 3 años y de 3 a 6 años, siendo antes de esta edad poco influyente el ambiente (Robinson, 1980). Como ésto no resuelve satisfactoriamente el tratamiento diferencial, dentro de cada grupo de edad, se deben contemplar diferencias individuales, y parece que el **rendimiento previo** (al menos respecto a objetivos de las áreas numérica y verbal) es el mejor predictor del aprendizaje infantil.

2) **PROGRAMA**: Incluía las variables Modelo de Intervención, Calidad, Tiempo, Objetivos y Mediación Instructiva. Uno de los elementos importantes en el primitivo esquema era el **Modelo de intervención** debido a que, en base a los estudios anteriores, diferentes modelos se relacionaban a diferentes resultados. La clasificación de la que partíamos era la utilizada en el proyecto Head Start (centrada en la *filosofía educativa* predominante), pero habíamos comprobado la dificultad de clasificar la intervención real en el aula siguiendo sus directrices. A pesar de que durante este tiempo la hemos encontrado validada empíricamente (Kagan y Smith, 1988; Rusher et al., 1992), en el contexto real se puede comprobar la cantidad de información que se pierde sobre la dinámica del aula. Lo mismo ocurre al utilizar el criterio de *grado de estructura* del programa (Prescott, 1991). Otro criterio, utilizado en los estudios ingleses (Osborn y Milbank, 1987) es el *tipo de institución* (Nursery, Playgroup, Nursery class y Childminder). Aunque es operativo en su definición, ocurre lo mismo que en los criterios anteriores, existe casi la misma heterogeneidad entre-centros que intra-centros, por lo que resulta poco funcional, además todos estos criterios no tienen en cuenta que el programa puede variar durante el proceso de intervención; es decir, que de la planificación a la puesta en práctica, el programa puede diferir sensiblemente. Es por esta razón que la evaluación de la *implementación* de un programa es un paso necesario en la evaluación de la EI (una posible forma de medir la implementación mediante «levels of use of an innovation (LoU)» (Hall y Loucks, 1977). Una aproximación más cercana al proceso de aprendizaje es la utilización del **profesor** en vez de analizar métodos o conductas. Los estudios consultados ponen énfasis en las conductas del profesor, influidas por su preparación y composición del grupo (Howes y Marx, 1992), o un paso más en la operacionalización de tales conductas se centraría en las **tareas** que los profesores piden a sus alumnos. Las tareas parecen el mejor indicador de las diferencias cualitativas de las experiencias de aprendizaje individual (Bennett et al. 1984).

Respecto al elemento **calidad**, intentamos en un primer momento conocer los indicadores de calidad de un programa de forma que esto nos informara de aquellas características necesarias que un programa debería proporcionar para que repercutiera positivamente en los resultados esperados. Hasta ahora habíamos descubierto variables *estructurales* (tamaño del grupo, ratio profesor-alumno, formación del profesor, planificación, e implicación de los padres) y variables *procedurales* (aquellas medidas en instrumentos como ECERS e ITERS). Sin embargo, parece que estos indicadores no son tan fijos a todos los programas como en un principio se pensó, sino que parece que pueden variar

según el grado de estructura de cada programa (Wolery y Bailey, 1984) por lo que lo más adecuado sería analizar indicadores diferenciales de calidad, controlando por tipo de programa, antes de determinar las características de una buena intervención. Otra alternativa podría ser definir calidad como *poder de adaptación*. Bajo este enfoque se han utilizado tres posibles vías según consideremos el grado de: a) adaptación programa-niño, en este caso se han seleccionado niños de distintos niveles de rendimiento y analizado, retrospectivamente, los componentes del programa que pudieron causar las diferencias. Este tipo de diseño tiene las limitaciones de todo estudio *ex-post-facto*. b) adaptación profesor-niño, no sólo se centran en el estudio de conductas eficaces sino que dan un paso más al categorizar las conductas del profesor y concluir que varían en cada profesor y que dependen de determinadas condiciones (Crahay y Delhaxhe, 1991) c) adaptación tarea-niño, a semejanza de los estudios de Bennett (et al., 1984; et Kell, 1989), en el que estudia la adaptación durante el proceso de aprendizaje.

Respecto a la variable **tiempo** tiene ya menos sentido al eliminar la variable modelo. No hacer más comentario que en el estudio de la eficacia infantil, toda intervención que analicemos debe referirse a programas regulares, practicados en el aula y no a intervenciones innovadoras de corta duración o sin continuidad.

Por último, dentro del factor programa, comentaremos juntas las variables **objetivos** y **mediación instructiva**, puesto que definimos mediación como el grado de estructura de la tarea o variaciones que hace el profesor en función de las características individuales del alumno para que logre el objetivo planificado. Parece que, como explica Bennett, la conexión entre profesor y alumnos durante el proceso de aprendizaje se operativiza en las tareas y, en consecuencia, éstas pudieran ser las que mejor informaran del proceso de aprendizaje realizado en el aula.

En resumen, parece coherente redefinir el factor programa como proceso instructivo y evaluar el grado de adaptación tarea-niño en los términos de Bennett (operacionalización de calidad). Así como eliminar la variable modelo de intervención y centrarse en el análisis de tareas ayudados por profesores que tengan experiencia en esta etapa educativa.

3) **PADRES**: a pesar de comprobar la influencia de los padres en el desarrollo y aprendizaje del niño preescolar decidimos (por la dificultad de entrar en contacto y controlar dicho factor) centrarnos en el marco interno de la escuela y, en consecuencia, eliminarlo del modelo por el momento.

4) **COMPETENCIA INDIVIDUAL**: ¿qué destrezas cognitivas y sociales debe tener un niño al finalizar la intervención o al finalizar la etapa preescolar? Una pregunta que parece tan básica no parece tener una única respuesta. Este problema está íntimamente ligado al del problema de la medida en esta etapa, pero según hemos visto pudiera tener una doble solución: 1) Basarse en los resultados de **múltiples medidas** (estandarizadas y observación de profesores) para definir al niño competente como hiciera Robinson (1980). Este enfoque pudiera ser válido si se optara por elaborar instrumentos que midieran aspectos normativos del desarrollo y aprendizaje infantil; si fuera así, lo más conveniente sería utilizar Índices de cambio (Rosenberg et al., 1987) para describir el impacto del programa en el niño. 2) Dar prioridad al currículum, evaluando la **respuesta emocional y cognitiva a las tareas** de clase. Si fuera esta la opción, se podría utilizar las Ratio de progreso y Ratio de consolidación propuestas por Bennett (et al., 1984).

Objetivo 2º: Adecuar la metodología de investigación, hemos comprobado que las limitaciones encontradas por nosotros son similares en otros contextos puesto que el enfoque experimental sigue dominando a la hora de diseñar los estudios sobre eficacia de la Educación Infantil. No obstante, en el estudio de la calidad educativa, actualmente, se está trabajando en otro enfoque que creemos puede enriquecer y ayudarnos a comprender mejor el aprendizaje infantil individual durante esta etapa educativa, y que podemos resumir en los siguientes puntos:

a) Ratificamos el abandono del paradigma input-output por otro que considere el proceso de aprendizaje en el aula. Es preciso analizar la realidad intentando comprender la complejidad del aula y, concretamente, cómo se adapta el profesor al alumno y viceversa y cómo se adaptan profesor y alumno al entorno de la clase. Además, cuando hablemos de calidad lo debemos hacer refiriéndonos al contexto (describiendo las características del entorno y sus recursos) y al contenido de la experiencia de aprendizaje (describiendo la tarea). Esto supone un cambio de paradigma de investigación. La utilización del «Model of classroom learning processes» (Bronfenbrenner, 1976; Doyle, 1979; y operativizado en Bennett et al., 1984; Bennett y Kell, 1989) parece ser hasta el momento uno de los mejores y más productivos modelos de investigación para conocer la forma de optimizar el aprendizaje preescolar.

b) Las variables independientes de la investigación en el campo de la EI deben ser seriamente consideradas definiéndolas, como generalmente se aconseja, operativa y claramente pero, además, se deberían incluir otros dos apartados: 1) Base teórica que sirva de guía a la investigación, y 2) Definición y Medida de la fidelidad del tratamiento (implementación) (Lelaurin y Wolery, 1992).

c) Respecto a la medida de la competencia infantil, pudieramos clasificarla según el tipo de información proporcionada y utilizar aquella que más se acerque a nuestras necesidades o a los objetivos de investigación: 1) Posición del niño en un grupo normativo: Medir mediante índices de cambio (ya explicado anteriormente); 2) Posición del niño en una secuencia de desarrollo: utilizar MAPS (Measurement and planning system, Bergan et al., 1991); 3) capacidad independiente de dificultad del ítem y del grupo: Rasch Measurement Model (Snyder y Sheehan, 1992); 4) Resultados individuales durante la tarea y a corto plazo: observación (Sylva et al., 1986) y entrevista (Bennett et al., 1984).

d) Respecto a los análisis de datos, dos pudieran ser las tendencias más útiles: 1) si se evalúan los programas de forma global, parece que la técnica de valor añadido es bastante adecuada puesto que nos permite evaluar los efectos del programa cuando no se ha podido seleccionar un grupo de control o no ha sido posible la manipulación del tratamiento (la mayoría de las veces en este tipo de estudios y que es una de las causas más frecuentes de invalidez de los resultados); 2) Si se pretende analizar la complejidad del aula se pueden combinar técnicas cuantitativas y cualitativas, si bien predominan estas últimas puesto que parecen ofrecer mejor comprensión de los datos en este tipo de modelos.

REFERENCIAS

- BENNET, N., DESFORGES, C., COCKBURN, A. y WILKINSON, B. (1984): *The quality of pupil learning experiences*. Londres: LEA.
- BENNET, N. y KELL, J. (1989): *A good start? Four year olds in infant schools*. Oxford: Basil Blackwell Ltd.
- BERGAN, J. R., SLADDECZEK, I. E. y SCHWARZ, R. D. (1991): Effects of a measurement and planning system in kindergarten. Cognitive development and educational programming. *American Educational Research Journal*, 28 (3), 683-714.
- BRONFENBRENNER, U. (1976): The experimental ecology of education. *Teachers College Record*, 78 (2), 157-204.
- CRAHAY, M. y DELHAXHE, A. (1991): How do preschool teachers requests influence children's behaviors? *Teaching & teacher education*, 7 (3), 221-239.
- DOYLE, W. (1979): Classroom tasks and studen abilities. En P. L. PETERSON y H. J. WALBERG (eds.), *Research on teaching: concepts, findings and implications*. Berkeley, Ca: McCutchan.
- HALL, G. E. y LOUCKS, S. F. (1977): A developmental model for detrmining whether the treatment is actually implemented. *American Educational Research Journal*, 14 (3), 263-276.

- HOWES, C. y MARX, E. (1992): Raising questions about improving the quality of Child Care: Child care in the United States and France. *Early Childhood Research Quarterly*, 7, 347-366.
- KAGAN, D. M. y SMITH, K. E. (1988): Beliefs and behaviours of kindergarten teachers. *Educational research*, 30 (1), 26-35.
- LELAURIN, K. y WOLERY, M. (1992): Research standards in Early intervention: defining, describing, and measuring the independent variable. *Journal of Early Intervention*, 16 (3), 275-287.
- OSBORN, A. F. y MILBANK, J. E. (1987): *The effects of early education*. Oxford: Clarendon Press.
- PRESCOTT, E. (1991): Approaches to quality in early childhood programs. En J. D. QUISENBERRY et al. (eds.): *Reading from childhood education*, Vol. II. Illinois: Association for childhood education interantional, 201-207.
- ROBINSON, P. A. (1980): Research and the child: the family. En D. G. RALE et al. (eds.): *Aspects of early childhood education. Theory to research to practice*. Nueva York: Academic Press, 119-142.
- ROSENBERG, S. A. et al. (1987): An empirical comparison of formulas evaluating early intervention program impact on development. *Exceptional children*, 54 (3), 213-219.
- RUSHER, A. S., MCGREVIN, C. Z. y LAMBIOTTE, J. G. (1992): Belief system of early childhood teachers and their principals regarding early childhood education. *Early Childhood quarterly*, 7 (2), 277-296.
- SNYDER, S. y SHEEHAN, R. (1992): The Rasch Measurement Model: an introduction. *Journal of Early Intervention*, 16 (1), 87-95.
- SYLVA, K., ROY, C. y PAINTER, M. (1986): *Childwatching at playgroup and nursery school*. Oxford Preschool Research Project. Grant McIntyre Ltd.
- WOLERY, M. y BAILEY, D. B. (1984): Alternatives to impact evaluations: suggestions for program evaluation in early intervention. *Journal of the Division for Early Childhood*, 9 (1), 27-37.

LA METODOLOGÍA MICROGENÉTICA: UNA ALTERNATIVA METODOLÓGICA EN INVESTIGACIÓN PSICOPEDAGÓGICA BÁSICA

Mercè Garcia-Milà

Facultat de Ciències de l' Educació. Universitat de Barcelona

INTRODUCCIÓN

Toda intervención psicopedagógica va dirigida a obtener un cambio cognitivo, el cual puede ser tanto de actitudes como de procedimientos. La Ciencia Cognitiva ha hecho grandes aportaciones al estudio de ambos tipos de cambio, sin embargo, las metodologías tradicionalmente utilizadas nos han dado poca información sobre dos aspectos fundamentales del cambio: el proceso mediante el cual tiene lugar el cambio, y las diferencias individuales que se observan durante tal proceso. La investigación psicopedagógica ya no se conforma con el estudio de los estados puntuales previos a una intervención y posteriores a ella, dando cuenta sólo de los estados cognitivos inicial y final del sujeto en estudio; ni con una formulación generalizada de dicho cambio, olvidando, por tanto las diferencias individuales. Los métodos clásicos de investigación utilizados para estudiar cambios producidos mediante la intervención psicopedagógica son, por un lado, el método longitudinal, el cual estudia el mismo sujeto en sucesivos estados distribuidos en un periodo de tiempo largo. Este método nos da información sobre estado puntual de indicadores cognitivos en estudio, estabilidad y resistencia a la intervención, a la vez que las diferencias individuales. En contraste, el método transversal estudia a grupos de sujetos representativos de los distintos estados cognitivos y nos da información generalizada sobre el efecto puntual de la intervención en su inducción al cambio cognitivo. Ninguno de los métodos citados proporciona información sobre el proceso generador del cambio cognitivo, objetivo fundamental de los estudios psicopedagógicos. El método microgenético permite alcanzar dicho objetivo.

EL MÉTODO MICROGENÉTICO

El método microgenético consiste en estudiar individualmente a un grupo de sujetos durante una serie de sesiones en las cuales se le presenta a cada uno una tarea cuya solución implica la aplicación de estrategias de aprendizaje o actitudes objeto de estudio. Las sesiones se extienden desde el comienzo del cambio hasta el momento en que se alcanza relativa estabilidad. Se graban las sesiones en video o audio, de forma que la conducta cognitiva en estudio pueda analizarse tanto cualitativamente (estudios de caso) como cuantitativamente (análisis estadístico). El material registrado es transcrito y codificado por dos experimentadores con el fin de determinar la fiabilidad y asegurar la objetividad. El

análisis estadístico se lleva a cabo a partir del material codificado mediante diseños intra-sujeto de medidas repetidas.

La alta densidad de observaciones durante el periodo de cambio cognitivo que se inicia con el descubrimiento de la estrategia hasta que su uso se hace estable y generalizado permite un análisis en profundidad de tal proceso, el cual proporciona información «sobre las etapas y circunstancias del cambio, el cambio *per se* y el proceso de generalización del cambio más allá del contexto inicial» (Siegler y Crowley, 1991, p. 608). Según estos autores, el análisis detallado de la conducta observada permite inferir los procesos que dan lugar tanto a los cambios cuantitativos tradicionalmente medidos mediante índices de rapidez y exactitud en las respuestas, como a los cambios cualitativos que acompañan a los anteriores. Por tanto, las características de dicha metodología son: a) la recogida de datos es individualizada, el experimentador trabaja individualmente con cada sujeto; b) las sesiones experimentales del experimentador-sujeto se repiten a lo largo de un determinado periodo de tiempo; c) el periodo de tiempo se concentra desde antes de la adquisición de la estrategia o actitud hasta que su uso se estabiliza; d) el sujeto trabaja con una o varias tareas, cuya solución pone de manifiesto su estado cognitivo respecto a la estrategia o actitud objeto de la intervención psicopedagógica; e) la alta densidad de observaciones que permite la repetición de la tarea a lo largo del proceso experimental de recogida de datos proporciona un análisis secuencial del proceso de cambio; y por último, f) el registro de protocolos permite llevar a cabo un análisis cualitativo del cambio cognitivo.

La metodología resulta ideal para proyectos de investigación cuyo objetivo es el estudio de la adquisición de una determinada habilidad o actitud, su aplicación de forma generalizada para obtener un cambio conceptual estable y su transferencia a nuevos campos de aplicación. Por tanto, todos aquéllos estudios de intervención psicopedagógica de generalización y transferencia cognitiva son especialmente adecuados para aplicar la metodología microgenética. Un aspecto fundamental de esta metodología es que permite poner de manifiesto la evolución metacognitiva del sujeto, cuán consciente éste es de la adquisición de habilidades o actitudes y de su eficacia para resolver tareas nuevas, y por tanto del proceso de generalización y transferencia anteriormente citados.

EJEMPLOS DE APLICACIÓN DEL MÉTODO

El uso del método microgenético se origina en estudios evolutivos del aprendizaje. El concepto nace de estudios de H. Werner, y Vygotsky (1978), citando a este autor, confirma la eficacia de un método que estudie el proceso de cambio, en lugar de estudiar procesos cognitivos cuyo desarrollo ya se encuentra completo y *fosilizado* (p. 65). Aunque el método microgenético no se usa de una forma generalizada, existen ya en la actualidad según Siegler y Crowley (1991) una veintena de estudios.

La mayoría de estudios publicados que utilizan la metodología microgenética son estudios cuyo objetivo es describir la adquisición y evolución de una determinada estrategia cognitiva. Algunos ejemplos pioneros son Karmiloff-Smith (1984), la cual analizó cambios en la representación mental en niños de cuatro a diez años en la tarea piagetiana de equilibrar la balanza; Siegler y Jenkins (1989) trabajaron con niños de cuatro y cinco años cuando descubren y generalizan la *estrategia del mínimo* (*min strategy*), estrategia aritmética de la suma que implica contar desde el sumando más grande tantas veces como indique el sumando pequeño. En otro campo de investigación aunque utilizando el mismo método, otros autores (Kuhn y Phelps, 1982) investigaron el desarrollo de las habilidades de experimentación en pre-adolescentes; Schauble (1989) estudió también las estrategias de experimentación en niños de ocho años en un sistema de atribución causal, haciendo hincapié en las diferencias individuales en un grupo de 22 sujetos. En la misma línea de investigación, Kuhn, Schauble y Garcia-Milà (1992) compararon el desarrollo de habilidades de razonamiento inductivo en un medio informatizado y uno real, sin encontrar diferencias en el desarrollo de dichas habilidades en los que se refiere al

medio de presentación de la tarea cuando el estudio se realizó individualizado, pero encontrando grandes diferencias individuales en el patrón de desarrollo común a ambos medios

El mismo tipo de estrategias de experimentación y habilidades de razonamiento inductivo se estudió en adultos (Kuhn, García-Milà, Zohar, sometido a publicación). El objetivo principal en dicho estudio fue la investigación de las propiedades generalidad y transferencia que presentan los procesos de razonamiento inductivo implícitos en la revisión de teorías en dominios distintos de conocimiento. El estudio puso de manifiesto la inadecuada presentación del método científico en los capítulos introductorios de Física y Química de Bachillerato. Los sujetos fueron 17 alumnos de primer curso universitario de la asignatura de Filosofía y Razonamiento Lógico, los cuales asistieron a dos grupos de diez sesiones paralelas (un total de 20 sesiones). Las tareas consistieron en una estructura multicausal, en la cual cinco variables presentaban un efecto en el sistema. Las variables se escogieron de forma que estuvieran relacionadas con las teorías previas de los sujetos. El objetivo de la tarea era el de descubrir la estructura causal del sistema mediante la experimentación. El análisis microgenético mostró que los sujetos identifican las relaciones causa-efecto a partir de un repertorio de inferencias tanto válidas como inválidas, sin ser conscientes de la eficacia de las inferencias válidas, ni de los problemas que conlleva el utilizar inferencias inválidas. La selección de la estrategia adecuada entre el repertorio del que disponían los sujetos mostró gran variabilidad, es por ello que una recogida de datos puntual hubiera dado una caracterización errónea. El análisis cualitativo que permitió el método microgenético facilitó identificar la coexistencia de estrategias válidas e inválidas, proporcionando un examen detallado de los patrones individuales de cambio.

A partir de los resultados, se pudo concluir que es difícil identificar un proceso de desarrollo único que caracterice a todos los sujetos. El desarrollo metaestratégico es el que parece motivar los procesos de transferencia según los datos cualitativos obtenidos. Las diferencias individuales encontradas permiten concluir que los estudios transversales aplicados a estudios de cambio cognitivo pueden dar lugar a resultados engañosos, ya que especialmente después de la adolescencia, las diferencias en la población dentro de un grupo de edad a menudo exceden las diferencias entre grupos. Las comparaciones basadas en estudios microgenéticos permiten analizar estudios de caso en función de la propia plasticidad.

CONCLUSIÓN

Hasta el momento, la mayoría de los estudios de aplicación del método microgenético son estudios evolutivos. Nuestra propuesta consiste en utilizar la metodología microgenética en investigaciones de cambio cognitivo inducidos por una intervención psicopedagógica dirigida tanto al cambio conceptual y/o estratégico, como al cambio de actitudes. Todos aquellos cambios no lineales que tienen lugar mediante sucesivos avances y retrocesos son particularmente apropiados para esta metodología. Un ejemplo representativo sería el estudio de la formación y eliminación de preconceptos en el aprendizaje. La metodología microgenética nos daría información sobre su estado, estabilidad y su resistencia a la instrucción. También nos informaría sobre las condiciones cognitivas que preceden a los sucesivos estados de progreso y regreso de una forma individualizada, ya que en una metodología de medidas repetidas, el sujeto actúa de propio control, permitiendo el estudio de las diferencias individuales en el aprendizaje.

Dentro del estudio de los cambios cognitivos, otros campos importantes de aplicación podrían ser las investigaciones que analizan los cambios cognitivos de actitudes, donde la información sobre la efectividad de programas de orientación podría ser evaluada a partir de las diferencias individuales de respuesta ilustrada mediante el análisis cualitativo, el cual permite trazar patrones individuales que se desvían del patrón del grupo, entendiendo la variabilidad y la consistencia en los patrones de cambio.

El método microgenético presenta varios problemas. La principal crítica se refiere a su uso para estudios evolutivos, en los cuales se pueden confundir los efectos experimentales asociados a la recogida de datos con el propio desarrollo evolutivo (Pressley, 1992). Otro problema es el coste, ya que requiere además de sesiones individualizadas experimentador-sujeto, la transcripción del material registrado, y su codificación, la cual debe ser realizada por dos codificadores con el fin de asegurar la objetividad de los resultados. Sin embargo, el análisis estadístico en conjunción con el análisis cualitativo (estudios de caso) proporciona una riqueza de datos que compensa el coste adicional. Por otro lado hay que tener en cuenta el desgaste en los sujetos, el cual se manifiesta mediante el alto grado de mortalidad experimental y absentismo. Este problema debe subsanarse incluyendo inicialmente más sujetos de los necesarios. Por último, según Pressley (1992) existe el riesgo de que el cambio objeto de investigación no se produzca durante el periodo de tiempo que dura la intervención y recogidas de datos. Este sería un problema en los estudios evolutivos, pero no en los estudios interventivos, donde la ausencia de cambio sería ya informativa en sí misma. En definitiva, en estudios en los que el cambio es el objetivo de estudio, la ventaja de obtener una secuencia de fotos instantáneas tomadas antes, después, y en intervalos frecuentes durante el cambio cognitivo en lugar de inferir la conducta a partir de una o varias fotografías justifica los riesgos y problemas planteados.

REFERENCIAS

- KARMILOFF-SMITH, A. (1984): Children problem solving. In A. L. BROWN, & B. ROGOFF (Eds.): *Advances in developmental psychology* (Vol. 3, pp. 39-90). Hillsdale, NJ: Erlbaum.
- KUHN, D., y PHELPS, E. (1982): The development of problem solving strategies. In H. REESE (Ed.): *Advances in child development and behavior* (pp. 1-44). New York: Academic Press.
- KUHN, D., SCHAUBLE, L., & GARCÍA-MILÀ, M. (1992): Cross-Domain development of scientific reasoning. *Cognition and Instruction*, 9(4), 285-327.
- KUHN, D., GARCÍA-MILÀ, M., & ZOHAR, A. (sometido a publicación): Microgeneteic analysis of inductive causal and noncausal inference. *Cognitive Psychology*.
- PRESSLEY, M. (1992): How not to study strategy discovery. *American Psychologist*, 47(10), 1240-41.
- SCHAUBLE, L. (1990): Belief revision in children: The role of prior knowledge and strategies for generating evidence. *Journal of Experimental Child Psychology*, 49, 31-57.
- SIEGLER, R. S., y CROWLEY, K. (1991): The microgenetic method. A direct means for studying cognitive development. *American Psychologist*, 46(6), 606-620.
- SIEGLER, R. S., y CROWLEY, K. (1992): Microgenetic methods revisited. *American Psychologist*, 47(10), 1241-43.
- SIEGLER, R. S., y JENKINS, E. (1989): *How children discover new strategies*. Hillsdale, NJ: Erlbaum.
- VYGOSTKY, L.S. (1978): *Mind and Society: The development of higher mental processes*. Cambridge, MA: Harvard University Press.

INTERVENCIÓN EDUCATIVA A TRAVÉS DE LA INVESTIGACIÓN COOPERATIVA

Leonor Buendía Eisman
Honorio Salmerón Pérez
Universidad de Granada

1. INTRODUCCIÓN

La actitud negativa de la mayoría de los maestros a realizar investigación educativa es de todos conocida, incluso después de realizar algún curso de formación (Pettus and Allain, 1991).

Waxman, Freiberg and Knight (1986) haciendo un recorrido por el contenido de cursos sobre investigación educativa, comentan que estos raramente ayudan a los profesores a entender y aplicar la investigación. Como consecuencia el desánimo y el desinterés es la tónica dominante cuando se intenta organizar un curso de investigación educativa dirigido a maestros en ejercicio.

En un reciente trabajo Sardo-Brown (1990), pone de manifiesto dos formas que él ha utilizado para motivar y trabajar con profesores en cursos sobre investigación educativa.

En una, los profesores se implicaron en un proceso de investigación-acción. En el segundo, les proporcionaron informes de investigación para ser leídos y comentados.

Los artículos leídos fueron:

a) Schunk, D. H. (1982): Progress self-monitoring: Effect of children's self-efficacy and achievement. **Journal of Educational Research**, 89-93

b) Wolfe, D. A., Mendes, M. G. & Factor, D. (1984): A parent administered program to reduce children's television viewing. **Journal of Applied Behavior Analysis**, 17, 267-272.

Después de trabajar estas investigaciones los profesores se consideraron importantes los tópicos que en esos materiales se trabajaban y en la mayoría de los casos eran sentidos como problemas que debían ser investigados, pero mostraron su desconfianza ante la posibilidad de generalizar a otras situaciones los resultados obtenidos. No obstante, las principales dificultades aparecieron a la hora de interpretar la información estadística contenida en los artículos a la vez que se consideró que el lenguaje utilizado en los informes estaba bastante lejos del que emplean los profesores para manifestar los problemas de la vida en la clase.

Por el contrario los que participaron en proyectos de investigación-acción manifestaron una mayor confianza en la investigación educativa e incluso su actitud hacia la posibilidad de aplicar los resultados de la investigación en la mejora de la práctica educativa fue mucho más positiva. Otros autores también abogan por cursos de formación desde la realización de proyectos de Investigación-Acción (Harwood, 1991; Hutchinson, 1991) para evitar deficiencias generadas con modelos diferentes (Stenhouse, 1975).

Nosotros, basándonos en los trabajos de Sardo-Brown (1990,1991), Shalaway (1990) y Sardo-

Brown (1992) organizamos un seminario permanente de investigación educativa a través de la participación de todos los profesores en un proyecto de Investigación cooperativa.

Consideramos que el seminario debía comenzar conociendo la actitud hacia la investigación educativa de los profesores que participaban en él y plantear una investigación cooperativa con todos ellos que permitiera una implicación en los procesos de cambio en las aulas de aquellos problemas que en las sesiones del seminario surgieron como más importantes.

Una vez realizada la investigación cooperativa volveríamos a medir la actitud para conocer la tendencia del cambio (si es que se producía) hacia la investigación educativa en general, y la necesidad sentida por los profesores.

Así pues, el objetivo específico de nuestro estudio fue: **«Conocer la actitud de los profesores hacia la investigación educativa y la posible tendencia del cambio de actitud una vez realizado un curso de formación y participación en un proyecto de investigación cooperativa».**

La investigación cooperativa es una modalidad de investigación-acción. Su definición —aunque con variantes en los distintos autores— pone el énfasis en el hecho de que investigadores y educadores (no se excluyen otros miembros de la comunidad educativa) trabajan juntos en la planificación, implementación y análisis de la investigación que se lleva a cabo para resolver problemas inmediatos y prácticos de los maestros, compartiendo la responsabilidad en la toma de decisiones y en la realización de las tareas de investigación, (Bartolomé y Anguera (Coord.), 1990: 31).

2. MÉTODO

2.1. Sujetos

El estudio fue realizado por los 24 profesores que participaron en el seminario de formación. Todos ellos eran profesores de primero de primaria, y por lo tanto habían comenzado los nuevos ciclos de la L.O.G.S.E. en este curso.

Todos, excepto uno que trabaja en un centro privado, eran maestros de escuelas públicas. Había 20 mujeres y 4 hombre y la edad media era de 32,3 años.

2.2. Instrumentos

Para conocer la actitud de los profesores hacia la investigación educativa, el primer día de seminario realizaron una escala de actitud tipo Likert con 20 ítems y 5 posibles respuestas a cada ítem¹. La puntuación mínima posible en la escala fue de 20 y la máxima 100. Para conocer la consistencia interna de la escala se utilizó el coeficiente alpha, obteniendo un índice de .81. La fiabilidad test-retest se realizó con un grupo de profesores de características similares en cuanto que también eran de primaria y con edad media similar (33,4 años) que estaban realizando un curso de metodología de la investigación en el último semestre del año anterior. El período de tiempo entre test-retest fue de seis meses y se obtuvo un coeficiente de 0.79.

El instrumento utilizado fue

1. La investigación educativa es importante para la enseñanza.
2. Yo encuentro que la investigación es bastante contradictoria con la práctica educativa.
3. Es difícil entender la investigación educativa.

1 1= Totalmente de acuerdo, 2= Acuerdo, 3= Indiferente, 4= Desacuerdo y 5= Total desacuerdo.

4. La investigación me aporta bastante información para mi área de trabajo.
5. La investigación educativa es bastante teórica.
6. Encuentro que la investigación tiene aplicación a lo que yo hago en mi clase.
7. Utilizo los resultados de la investigación en mis clases prácticas.
8. Es importante para mí, como profesor obtener información sobre resultados de la investigación.
9. Realizar investigación por uno mismo, es una excelente idea.
10. Básicamente, bastante investigación educativa es utilizada para las clases.
11. Los informes de investigación están hechos con términos que son difíciles de entender.
12. Me gustaría trabajar con uno o dos profesores en un proyecto de investigación.
13. Me considero uno «consumidor» de informes de investigación.
14. Me creo capacitado para encontrar respuesta a los problemas que se plantean en mi clase.
15. Realizar seminarios o cursos de investigación educativa es una atractiva idea.
16. Adquirir conocimientos sobre cómo realizar una investigación es importante para mí.
17. Me gustaría aprender métodos de investigación participando en programas de entrenamiento.
18. Conocer los resultados de la investigación hecha en mi campo, es importante para mí.
19. Leer investigación es tiempo aprovechado.
20. Conociendo la investigación educativa sería un mejor profesor.

2.3. Procedimiento seguido

Realizada la encuesta se realizó una sesión de discusión de los problemas que más preocupaban en ese momento a los profesores que participaban en el seminario

De todos los temas tratados, el que más preocupaba, con vistas a los nuevos diseños curriculares, fue el de organización del aprendizaje y formación de los profesores en aprendizaje cooperativo. Sobre este tópico comenzamos el seminario. La duración fue de 40 horas, desde primeros de octubre hasta el 15 de diciembre. Las sesiones tuvieron lugar los martes y jueves de 7 a 9 de la tarde.

Diseñamos conjuntamente la investigación sobre «aprendizaje cooperativo» con el convencimiento de la mejora que supone para los resultados académicos y sociales de los alumnos, que el profesor utilice esta forma de organización del aprendizaje. La mayoría de las veces el aprendizaje cooperativo se utiliza en el aula de forma complementaria con el aprendizaje tradicional sin embargo Slavin (1983, 1984, 1986, 1987) Gutiérrez y Slavin (1992), Webb (1984) entre otros, señalan la superioridad de este tipo de aprendizaje en las diferentes materias, tanto en la consecución de logros sociales como cognitivos.

En una entrevista realizada a Daviol Johnson (1987) comenta que hay más evidencia validada de las importantes repercusiones del aprendizaje cooperativo que de ningún otro tópico educativo.

A partir de esta decisión, se dividieron los profesores del seminario en cuatro grupos

Cada uno se comprometió a trabajar en su aula con un modelo de aprendizaje cooperativo diferente, aunque todos pretendían cubrir los mismos objetivos (el tercer proyecto de matemáticas que correspondía a la adquisición del sistema de numeración decimal). Los grupos fueron:

- 6 profesores trabajaron con la técnica Jigsaw.
- 6 profesores siguieron la técnica Learning together (colaborativa).
- 6 Trabajaron con pequeños grupos (de 3 o 4 alumnos) sin meta común.
- 6 Siguieron trabajando con aprendizaje individual y en asamblea de clase.

Cada día, durante la primera hora se discutían los problemas que encontraban los profesores en la organización de los grupos en las clases, en la preparación del material y en el desarrollo del modelo.

Durante la segunda hora se desarrollaba un tema del seminario (sobre aprendizaje cooperativo) dirigido por nosotros pero preparado por ellos. Cada grupo se encargó de leer y exponer trabajos realizados sobre el modelo que ellos seguían en sus aulas.

Terminado el seminario y para conocer el nivel de logro de los objetivos que cada profesor se había propuesto, elaboramos una prueba criterial sobre los contenidos impartidos.

Las etapas que seguimos en la elaboración de las pruebas fueron:

a) Análisis crítico por parte de los profesores de los ítems elaborados individualmente.

b) Selección de ítems en función de diferentes criterios.

1. Coherencia entre el contenido del ítem y el objetivo que se pretende medir.

2. Coherencia con el desarrollo de las clases.

3. Importancia concedida por los profesores a las unidades temáticas y su reflejo en los ítems de la prueba.

4. Elaboración de una tabla de especificaciones de cada ítem con el contenido, proceso cognitivo que se pretende conseguir. Los totales de dicha tabla y su expresión en porcentajes, muestra la importancia concedida por los profesores a los diferente contenidos.

Cada profesor pasó sus correspondientes pruebas obteniendo el nivel de logro de cada objetivo por cada alumno y el nivel de logro del grupo también en cada objetivo. Igualmente el profesor consiguió una puntuación global de cada alumno en la prueba total. El profesor, dispuso de un gráfico por alumno que fue el que permitió diseñar su próximo proyecto docente y organizar nuevos grupos en función de los resultados de las diferentes pruebas criterioales.

2.4. Obtención de resultados

Puesto que nuestro objetivo era conocer la tendencia de cambio de actitud de los profesores hacia la investigación educativa, una vez terminada la experiencia, en el mes de febrero y una vez concluido el seminario y realizada la prueba criterial, volvieron a realizar la encuesta los 24 profesores.

Los resultados fueron:

	Punt.X en Pre.	SD	Punt.X en Post.	SD
1	3.83	0.61	2.08	0.61
2	2.17	0.37	2.17	0.37
3	2.17	0.37	2.33	0.71
4	3.27	0.48	2.33	0.62
5	1.82	0.75	3.26	0.30
6	3.21	0.81	1.17	0.76
7	4.11	0.58	1.85	0.92
8	1.66	0.37	1.78	0.41
9	2.87	0.97	1.17	0.61
10	3.17	0.70	1.17	0.80
11	2.08	0.56	3.18	1.32
12	1.68	0.78	1.08	0.86
13	2.17	0.37	2.96	0.93
14	1.82	0.38	4.25	0.60
15	2.17	0.61	3.83	0.76
16	2.17	0.78	1.17	0.35
17	3.83	0.88	1.82	0.60
18	1.82	0.38	1.82	0.38
19	4.16	0.21	2.84	0.30
20	4.16	0.21	1.17	0.61

Los ítems que indican la actitud hacia el uso de la investigación educativa han mostrado una clara tendencia de cambio. Por ejemplo los ítems 5 y 10, en el pretest tienen un valor medio de 3.21 y 3.17 respectivamente. Los valores en el postest cambiaron considerablemente obteniéndose en ambos un valor medio de 1.17, que indica un acuerdo total con la posibilidad de utilizar los resultados de la investigación educativa.

Otro conjunto de ítems está dirigido a descubrir las dificultades específicas que los profesores tienen cuando interpretan los resultados de la investigación.

Las cuestiones que trataron este uso cambiaron de diferente manera en la dirección deseada. Por ejemplo, el ítems 2 y 3 permanece con el mismo valor medio, sin embargo el 11 experimenta un pequeño cambio.

Los ítems (15 y 17) sobre necesidades formativas en investigación educativa variaron en mayor medida. En el ítem 17 desde una posición de desacuerdo con la sentencia «deseos de apuntar métodos de investigación participando en programas de entrenamiento» se pasa a un total acuerdo, una vez realizado el curso. Evolución similar aunque menos sobresaliente ocurrió en el ítem 15.

Varios ítems fueron diseñados para evaluar la actitud de los profesores hacia su propia investigación (9, 12 y 16). El mayor cambio se produjo en el ítem 16 cuyo valor medio fue de 1.17.

El ítem con mayor cambio en un sentido no deseado fue el 14, cuando los profesores conocieron las posibilidades de mejorar su práctica a través de la investigación educativa, la respuesta al ítem «me creo capacitado para encontrar respuestas a los problemas que se plantean en mi clase» cambió en sentido contrario manifestándose con una media de 4.25 o sea de «gran desacuerdo con el ítem».

Otros ítems igualmente importantes por los resultados, son el 5 y el 7. En el ítem 5: «la investigación educativa es bastante teórica» hubo bastante acuerdo en el pretest cambiando a un desacuerdo (valor medio 3.26) en el postest.

En el ítem 7, en el pretest hubo un casi total desacuerdo con el ítem pasando a tener un valor de 1.85 en el postest.

Aunque no hemos hecho diferencias significativas, porque solo nos interesa ver la dirección del cambio, en el ítem 20 que es un ítem globalizador del objetivo que nos propusimos, obtuvimos un valor t de 2.87 significativo al 0.05.

Concluyendo, los profesores que realizaron el curso de formación en aprendizaje cooperativo, implicándose a lo largo del mismo en un proyecto de investigación, dirigido al cambio metodológico en las aulas, han cambiado sensiblemente su actitud hacia la investigación educativa, considerándola importante para los procesos de innovación y mejoramiento de la práctica, a la vez que mayoritariamente comunicaron su satisfacción por el programa desarrollado en este curso de formación.

REFERENCIAS BIBLIOGRÁFICAS

- BARTLLET, L. (1991): The dialectic between Theory and Method in Critical interpretative Research. *British Educational Research Journal*, 17 (1), 19-34.
- BARTOLOMÉ, M. y ANGUERA, M^a. T. (1990): *La investigación cooperativa: vía para la innovación en la Universidad*. Barcelona: PPU
- BOHLMAYER, E. M. y BURKER, J. P. (1987): Selecting cooperative learning techniques: a consultative strategi guide. *School Psychology Review*. 16 (1), 36-49.
- BRANDT, K. (1987): On coepetation in schools: A conversation with David and Roger Johnson. *Educational Leadership*. Novembre, 14-18.
- GUTIÉRREZ, R. y SLAVIN, R. (1992): Achievement effects of the Nongraded Elementary School: A Best Evidence Systhesis. *Review of Educational Research*, 62 (4), 333-376.

- HARWOOD, D. (1991): Action Research us Interaction Analysis: a time for reconciliation? A Reply to Barry Hutchinson. *British Educational Research Journal*, 17 (1), 67-72.
- HUTCHINSON, B. (1991): Active tutorial work discussion and educational Research. *British Educational Research Journal*, 17 (1), 51-65.
- McEWEN, N. (1992): Luality Criteria for maximizing the use of Research. *Educational Resercher*, 21 (7), 20-22 y 27-32.
- PETTUS, A. M. y ALLAIN, V. A. (1991): *Teachers and research: Thresholds of change*. Paper presented at the annual meeting of the Association of Teacher Educators. New Orleans.
- SARDO-BROWN, D. (1990): Middle level teachers perceptions of action research. *Middle School Journal*, 22 (2), 30-32.
- SARDO-BROWN, D. (1991): Secondary teachers participation in action research. *The High School Journal*, 75 (1), 48-58.
- SLAVIN, R. E. (1983): *Cooperative Learning*. Longman. New York.
- SLAVIN, R. E. (1987): Cooperative learning and the cooperative school. *Educational Leadersip*. Novembre, 7-13.
- SLAVIN, R. E. (1990): Achievement effects of Ability grouping in Secondary Schools. A Best-evidence synthesis. *Review of Educational Research*, 60 (3), 471-499.
- STENHOUSE, L. (1975): *an introduction to curriculum Research and Development*. Heinemann. London.
- WAXMAN, H. C.; FREIBERG, J. y KNIGHT, S. (1986): *Using research knowledge to improve teacher education: Teachers perceptions of the value of educational research*. Paper presented at the annual meeting of the Association of Teacher Educators. Atlanta, GA.
- WEBB, N. (1984): Interacción entre estudiantes y aprendizaje en grupos pequeños. *Infancia y Aprendizaje*, 27/28, 159-183.

LA INVESTIGACIÓN-ACCIÓN COMO MODELO DE INTERVENCIÓN PARA EL DESARROLLO DE LA AUTOGESTIÓN Y LA AUTONOMÍA PERSONAL

Isabel López, Soledad Romero, Teresa Padilla, Miguel Vargas, Dolores Martín
Universidad de Sevilla

Dedicamos este trabajo a Pilar Vázquez Labourdette, por todos los esquemas que nos ha ayudado a romper.

1. INTRODUCCIÓN

El objetivo de esta comunicación es realizar un primer intento de sistematización de un conjunto de experiencias de investigación-acción en las que estamos participando desde el curso escolar 1990/91. Dichas experiencias, que abarcan varios campos de intervención psicopedagógica, son las siguientes:

* La investigación-acción como metodología de formación y teorización del profesor de educación de adultos desde su práctica.

* Proyecto Horizonte: la investigación-acción como base de una experiencia de educación vocacional.

* Proyectos de innovación de la Enseñanza Universitaria.

* Proyecto de Formación de Padres.

* Proyecto de innovación y experimentación «Rincones y talleres: un modelo organizativo en educación primaria».

* Programa de exploración y estudios pedagógicos «Europa y la ética del sujeto ciudadano» (Oficina Franco-Alemán para la Juventud, O.F.A.J.).

* Proyecto de investigación-acción y formación permanente en el centro.

Vamos a centrarnos en el análisis de los procesos de **autonomización** y **autogestión** de individuos y grupos, como procesos que se dan en la medida en que se desarrolla la investigación-acción.

En torno a estos procesos (autogestión y autonomía), nos proponemos abrir una serie de cuestiones que nos planteamos como base para realizar nuestro análisis. Nuestra pretensión es esbozar, con la brevedad que requiere este tipo de trabajos, líneas de reflexión que se abren a partir de estos cuestionamientos. ¿Cómo se forma el grupo? ¿llega a existir el grupo en su más pleno sentido? ¿hasta dónde

un grupo tiene la capacidad de autogestión y autodeterminación? ¿qué papel juega el «animador» en todo este proceso? ¿pueden existir distintos niveles de autogestión? ¿qué relación existe entre el proceso de autogestión del grupo y la autonomía de la persona?

2. GENERACIÓN DE PROCESOS DE AUTONOMÍA Y AUTOGESTIÓN

El grupo es la base fundamental de la investigación-acción. Simultáneamente, a través de un proceso dialéctico, el colectivo va convirtiéndose en gestor de su propio desarrollo como grupo y como seres individuales. Es decir, si el grupo es imprescindible para la investigación-acción, ésta se convierte, a su vez, en base para que el grupo y sus miembros desarrollen su capacidad de autogestión y autonomía.

«No existe investigación-acción sin participación colectiva. (...) El desarrollo colectivo supone necesariamente el crecimiento del ser humano, tanto en el aspecto individual como colectivo» (Barbier, 1986).

La investigación acción es un medio que ayuda a generar un proceso de constitución de la persona y del propio grupo, proceso dialéctico en el que, en la medida en que el individuo se cuestiona y toma decisiones, también el grupo se ve movido a cuestionarse a sí mismo y tomar postura, lo que se vuelve de nuevo hacia el individuo, que se plantea nuevas reflexiones sobre sí mismo y su posición en el grupo y así sucesivamente. Cada uno de estos planteamientos sucesivos del grupo y del individuo pueden tener una doble consecuencia: o bien se refuerzan individuo y/o grupo, haciéndose más fuerte su implicación, o bien el grupo y/o el individuo abandona el proyecto que motivó el proceso.

«...En la reunión de gran grupo estábamos tratando de tomar ¡por fin! decisiones sobre las propuestas realizadas en la sesión anterior. Después de pasar una hora intentando tomar decisiones y no llegándose, como nos está pasando de forma continuada, **** se levantó y nos preguntó uno a uno y señalándonos con el dedo, si estábamos de acuerdo en votar. Se ve que no podía más con esta situación, se levantó y nos hizo tomar una decisión. Creo que este posicionamiento que tomó *** nos obligó a tomar una postura públicamente ante todo el grupo» (Diario de participante en el proyecto de OFAJ).

Los procesos de autogestión y autonomía

El grupo para poder entrar en procesos de autogestión, va pasando por diferentes fases:

1.— Se da en primer lugar una fase de contacto personal de los miembros del colectivo que se aglutina en torno a un proyecto, idea, necesidad y objetivo. Es el momento en que cada uno de los miembros del grupo trata de ir tomando su espacio dentro de éste.

2.— Superado este contacto inicial el poder que tiene el animador ha de crear estrategias de dinamización del grupo que permitan un clima de reflexión, de relajación, que despierte en cada uno de los miembros un espíritu crítico, solidaridad, colaboración, intercambio, cuestionamiento personal en todos los niveles y esferas de su vida.

«Para mi la experiencia está incidiendo en muchos aspectos de mi vida. No solamente en el trabajo, sino también en mis relaciones con mi pareja, con mis amigos, con mi familia y con mi entorno ambiental.(...) Mi forma de entender la vida no es la verla pasar y pasar

por ella, sino la de vivirla participando, cambiando y transformando realidades, que aparentemente eran inamovibles» (Diario de un participante del proyecto de formación de profesores de adultos).

Cuando el grupo se está constituyendo se dan intercambios de tipo personal que hacen entrar en juego dimensiones de carácter emocional que en un momento determinado pueden llevar a una situación de conflicto. Este conflicto también puede ser provocado de forma intencional por parte del animador, haciendo ver a los miembros del grupo que los problemas, la realidad que viven no es tan simple como ellos plantean inicialmente, sino que es más compleja. Ello lleva a los miembros del grupo a cuestionarse a sí mismos, cuestionamiento que implica todo un proceso de luchas internas, de necesidad de ser escuchado, de rupturas de esquemas.

«Recuerdo a aquella 'loca' profesora que siendo de 'organización' escolar vino a 'desorganizarnos' el tinglado (...). Primero nos dice que elaboremos entre todos el programa de la asignatura. 'Pero , ¿de qué va esta mujer?'. Y nosotros...¡pumba!, (...) a buscar contenidos para hacer el programa: primero los objetivos, después los contenidos 'contenidos'. Y ella: 'saber ser', además de 'saber hacer' y 'saber'. Pero...¿qué? ¿En la Facultad vamos a aprender a ser? ¡No, no, no!, Aquí venimos a aprender a ser buenos profesionales. Y ella: 'Sí, pero ¿qué es para vosotros 'ser buenos profesionales?'. ¡Y no te deja hasta que estás hecho polvo y no sabes por dónde salir. Y te cuestionas ¿y yo, qué hago aquí? ¿qué puedo hacer yo en todo este lío?» (Diario de una alumna de Pilar Vázquez en la Universidad).

El conflicto puede ser abordado por el grupo de dos formas: o bien decide reflexionar sobre el mismo para comprender la lógica de su funcionamiento; o bien se puede dar el proceso contrario, que no sea abordado y se de una situación traumatizante que rompa el grupo antes de que éste pueda llegar a una cierta madurez.

«Estamos estancados, creo que hay gente que se siente mal. Unos quieren que sigamos con sesiones reflexivas, otros quieren pasar ya a la acción. Pero ¿qué acción? Dicen que no tenemos nada y sí lo tenemos, pero en nuestros diarios. ¿por qué no trabajamos lo que tenemos en ellos? Creo que si lo analizásemos veríamos cuál es nuestro proyecto y podríamos seguir caminando. Si no lo hacemos esto se va al garete» (Diario de un participante del proyecto «Horizonte»).

3.— En un momento posterior, la superación del conflicto por cada uno de los miembros del grupo y de éste como colectivo, llega a través de la racionalización del mismo, de un proceso de reflexión e interiorización de todo lo ocurrido. Asimilado el conflicto como responsabilidad propia y del grupo, cada uno de los miembros del mismo se posiciona de una forma mucho más autónoma y libre respecto a sí mismo y al grupo y es el momento en el que el grupo está capacitado para implicarse en una mayor autogestión. Cuando cada uno de los miembros del grupo tiene su propia idea de la realidad que allí se está dando, es posible crear un posicionamiento de grupo, que toma protagonismo como tal, obligando a un reposicionamiento del animador. Este proceso va facilitando que cada persona vaya encontrando su propio espacio, emocional, sensitivo, racional, al interior del grupo de tal forma que pueda establecer un tipo de interacciones que vayan creando un clima de reflexión y de trabajo, de producción real y de expansión personal y colectiva.

«Actualmente me encuentro en un momento de fin y de principio, de cansancio y de fuerzas. De fin con respecto a la dependencia de la profesora y de principio con respecto a mi propia libertad de acción. Y en ese vaivén de polos enfrentados, parece más exacto decir que me encuentro en un momento de moratoria» (Diario de un alumno universitario).

En este momento no es que se haya conseguido la autogestión del grupo, sino que cada uno de sus miembros toma un mayor protagonismo, lo que conlleva a un reposicionamiento de todos los implicados, incluido el animador, produciéndose un renovación de roles que abrirá un nuevo bucle en la espiral del proceso de constitución y desarrollo del grupo.

Como síntesis, podemos entender la autonomía y autogestión de individuos y grupos como un proceso continuo, más que como un estado; como una situación perfeccionable, más que como un ideal estático.

«Hay que entender la autogestión (...) como una etapa y no como un estado (...) un proceso dinámico y temporal, histórico, de una conquista interminable pero progresiva» (Sánchez, 1983:47).

BIBLIOGRAFÍA

- AA.VV. (1991): «Trabajando con un grupo: dudas y reflexiones». *Comunicación presentada al II Congreso de Investigación Educativa de Estudiantes*, Sevilla.
- AA.VV. (1992): «La formación del profesorado de Educación de Adultos a través de la investigación-acción: la observación como proceso de formación». *Comunicación presentada al I Congreso de Investigación en Andalucía y Europa sobre Educación de Adultos, Movimientos Sociales y Universidad*. Sevilla.
- BARBIER, R. (1986): *Naturaleza de la Investigación-acción existencial. Documento interno*. Centro de Investigación sobre lo imaginario. Université Paris VIII.
- JURADO, C. y otros (1993): *Aprender a crear, crear para aprender*, Sevilla: Consejería De Educación y Ciencia de la Junta de Andalucía.
- LÓPEZ, I. y otros (1992): «Proyecto de formación de profesores de adultos desde la perspectiva de la investigación-acción». *Comunicación presentada al I Congreso de Investigación en Andalucía y Europa sobre Educación de Adultos, Movimientos Sociales y Universidad*. Sevilla.
- MARTÍN, D., PADILLA, T. y REBOLLO, M. A. (1991): «El modelo de investigación-acción como propuesta de análisis del proceso de formación-investigación de una escuela de padres». *Comunicación presentada al II Congreso de Investigación Educativa de Estudiantes*, Sevilla.
- SÁNCHEZ (1983): *Diccionario de Ciencias de la Educación*. Madrid: Santillana.

LA FORMACIÓN DE MADRES DESDE LA INVESTIGACIÓN-ACCIÓN

Teresa Padilla Carmona, M^a Dolores Martín Benítez
Universidad de Sevilla

INTRODUCCIÓN

Ser padres en el mundo actual es una tarea que produce angustia e incertidumbre. Muchas personas jóvenes, de una edad comprendida entre los 30 y los 40 años de edad, sienten la necesidad de ser formadas para poder ofrecer a sus hijos una educación adecuada. Solicitan conocimientos rudimentarios de psicología, pedagogía e incluso sociología que le sean útiles para abordar la tarea educativa de sus hijos, tarea que consideran de gran responsabilidad.

A ello hemos de añadir que dicha responsabilidad recae demasiado a menudo sobre la mujer, razón por la cual, existe un alto número de mujeres que demandan formación para la labor que les toca desempeñar, al mismo tiempo que buscan descargar en las conversaciones con otras compañeras, sus angustias, miedos, experiencias, problemas, etc.

Con esta breve introducción queremos poner de manifiesto la necesidad de emprender una acción educativa dirigida a estas mujeres que, a la vez de promover su formación para el rol de madres, promueva un análisis de los roles que la sociedad les encomienda desempeñar para así poder conseguir cambios y transformaciones en su realidad cotidiana.

Esta comunicación presenta una experiencia de formación de madres que se viene desarrollando desde el curso 1990-91 en el Centro de Educación de Adultos «José Sánchez Rosa», de la barriada de Pino Montano, situada en la periferia de la ciudad de Sevilla.

BREVE VISIÓN HISTÓRICA DEL PROYECTO DE ESCUELA DE MADRES

En este apartado nos parece oportuno aportar una breve referencia al trabajo desarrollado a lo largo de los cursos anteriores y al momento en el que actualmente nos encontramos.

En años anteriores (Martín, Padilla, Rebollo, 1991; 1992), las perspectivas del proyecto de ESCUELA DE MADRES giraban en torno al objetivo de conseguir que sus miembros llevaran a cabo proyectos y actividades conjuntamente con las APAs de los colegios públicos de la zona. Sin embargo, los procesos investigativos en la acción que se desarrollaron, pusieron de manifiesto que las necesidades de los miembros del grupo no coincidían con la finalidad marcada y que dichos miembros demandaban en su lugar una formación base que les permitiera partir de su situación concreta (su problemática familiar) hacia otra de carácter más general (el trabajo con colectivos como las APAs).

El rechazo de las madres hacia el proyecto fue tal que durante el curso 1990-1991, el grupo de personas en formación no avanzó hacia una línea común: mientras las mujeres insistían en formarse para la educación de sus hijos, los coordinadores intentaban realizar una labor más propia del Desarrollo Comunitario. De este modo, este curso se caracterizó por una dinámica en la que nadie estaba de acuerdo con el trabajo que se realizaba y por una profunda insatisfacción de todos sus miembros.

Dada la situación y de acuerdo con las características del procedimiento de investigación-acción, las dinamizadoras del grupo recogieron información sobre esa realidad problemática a través de un diario de campo y de entrevistas personales a todo los miembros. Cuando la información fue lo suficientemente completa se entregó al grupo, a la totalidad de mujeres que lo componían, y se dedicaron sesiones exclusivamente al análisis de lo que había sucedido, por qué se había llegado a ese punto y qué elementos habían entrado en juego.

Fue posible, en este momento, crear un clima más positivo en el que todos los miembros se mostraron de acuerdo con realizar su aportación para que, a partir del curso próximo, el proyecto de Escuela de Madres recogiera los intereses de todos sus miembros y se desarrollara de un modo más satisfactorio para todo el grupo. Por ello, el grupo definió y planificó el proyecto que quería: qué objetivos perseguía, qué metodología, qué posibles contenidos, etc. De modo que lo que actualmente se lleva a cabo dentro de la Escuela de Madres no es sino lo que las mujeres han planificado y planifican cada comienzo del curso escolar.

DESARROLLO METODOLÓGICO DE ESTA EXPERIENCIA DE FORMACIÓN

A la hora de comentar la metodología de formación que ponemos en práctica, queremos diferenciar dos procesos que se dan de una forma conjunta y paralela a lo largo de los encuentros del grupo. De este modo, en primer lugar, expondremos cómo se desarrollan los distintos temas de formación para pasar después a describir el proceso de reflexión/evaluación continua del aprendizaje y el grupo.

Desarrollo de los temas de formación

Las necesidades de formación y los objetivos de aprendizaje que este grupo persigue, si bien son planificados a principio de cada curso escolar, constituyen un proceso de revisión continua, de modo que los objetivos del grupo son matizados o modificados según la dinámica de trabajo que se va generando y según ésta incida en el grupo planteando en él nuevas necesidades de formación. De cualquier modo, podemos citar, entre otros, objetivos como los siguientes: Analizar la relación padres-hijos y la propia actuación educativa de los que componemos el grupo; Analizar la influencia de los medios de comunicación de masas en la educación informal y el desarrollo de niños y adolescentes; Estudiar y analizar el ambiente global en el que viven nuestros hijos: la barriada, las pandillas, los problemas de drogadicción,... buscando estrategias apropiadas para contrarrestar las influencias negativas detectadas; etc.

Estos objetivos van dando lugar a una serie de temas de formación (la Psicología Infantil; la figura de los padres en el proceso educativo de los hijos; relaciones padres-hijos; etc.), que son trabajados por el grupo desde una metodología que demanda de la **participación activa** de las mujeres. Dicha participación no la entendemos exclusivamente como la expresión de opiniones al grupo, sino también como compromiso con el diseño de cada tema, como aporte de información (tanto teórica como de casos prácticos) y como implicación en la evaluación de aprendizaje, del grupo y de uno mismo. De este modo, los temas son trabajados desde una doble perspectiva que intenta articular, por un lado, lo teórico (información/formación en torno a los temas) y, por otro, lo práctico (análisis de casos, observaciones de comportamientos en la familia,...).

Desarrollo del proceso de reflexión y crítica sobre el aprendizaje y el grupo

Como ya hemos mencionado, la metodología que usamos en esta experiencia se basa en el proceso de investigación en la acción y se caracteriza por ser flexible, abierta y emergente. Basándonos en las propuestas de diversos autores (Ander-Egg, 1990; Kemmis y McTaggart, 1988; López de Ceballos, 1987), hemos ido generando nuestro propio proceso de formación/investigación en un intento de conseguir los objetivos que este grupo se propone.

Un elemento inseparable de este proceso lo constituye la observación participante, de modo que, en el transcurso de las diferentes sesiones de encuentro, las dinamizadoras observan y toman nota de cuanto acontece en él, de modo que al grupo le sea posible realizar una reflexión sobre lo que hace y consigue. El análisis basado en estas notas de campo es lo que permite ir generando un proceso formativo adaptado a las necesidades del grupo. Por otro lado, desarrolla la capacidad (y actitud) de análisis en las mujeres del grupo, ampliando su habitual visión de los problemas educativos y relacionales, basada en concretos y particulares.

Por lo tanto, se intercala el trabajo cotidiano con sesiones en las que el grupo, con toda la información por delante, evalúa lo que se está llevando a cabo, aportando sugerencias, posibles cambios y modificaciones. Estas sesiones pueden ser concertadas, es decir, después de cada tema, se convoca una evaluación del tratamiento que el mismo ha recibido, o pueden ser fortuitas, esto es, se convocan porque surge una situación problemática o porque el grupo de madres así lo demanda.

REFLEXIONES FINALES

En este apartado nos gustaría poner de manifiesto que, en los distintos cursos que hemos desarrollado esta experiencia, no hemos encontrado con una circunstancia que imposibilita llevar a cabo una labor a largo plazo. Nos referimos al hecho de que cada año vienen mujeres nuevas y no permanecen siempre las del curso anterior (esta circunstancia suele darse a menudo en las experiencias de Educación de Adultos). Por ello y aunque puede parecer una visión un tanto pesimista, sentimos que estamos siempre en el mismo punto, que nunca avanzamos, que todavía nos hemos conseguido un grupo estable que profundice en su formación como madre y que, a partir de ahí empiece un cambio de roles, de actitudes y de esquemas mucho más complejo que la mera formación para la educación de los hijos. Si consiguiéramos un grupo estable, podríamos empezar trabajando las necesidades que denotan las mujeres pero también podríamos ir creándoles otro tipo de necesidades que actualmente no vivencian, necesidades como participar en los colegios de la zona en acciones dirigidas a la mejora de la calidad de educación y de vida de la misma. Este tipo de trabajo pertenece más al ámbito del Desarrollo Comunitario, pero pensamos que si pudiéramos conseguir estabilidad en el grupo de mujeres que asisten a la plataforma de Escuela de Madres, esta plataforma se nos mostraría como un campo muy fértil para este tipo de trabajos.

Por último, queremos señalar que el trabajo con estas mujeres es gratificante pero, a la vez, difícil. Provocar la reflexión y el análisis general y menos centrado en casos particulares es un proceso que ha de asumir el dinamizador y que exige que éste desarrolle en sí mismo unas capacidades y actitudes como la «escucha» del otro, la capacidad de sistematizar la vivencia del grupo, de dinamizar situaciones de conflicto, de provocar este conflicto en los demás y en sí mismo, etc. Quien, abandonando su rol de profesor tradicional, pretenda acometer esta labor, ha de ser consciente de que se encontrará en situaciones de incertidumbre y de cuestionamiento (incluso de sí mismo) y habrá de potenciar su propia formación en esta experiencia. Su capacidad de aprender día a día y de responder de forma creativa a los retos que se le presenten, serán aspectos esenciales que tendrá que poner en juego. A cambio, encontrará satisfacción y autorrealización en su trabajo, así como una mayor autonomía y desarrollo personal y profesional.

BIBLIOGRAFÍA

- ANDER-EGG, E. (1990): *Repensando la investigación-acción participativa*. Bilbao: Servicio central de publicaciones del gobierno vasco.
- LÓPEZ DE CEBALLOS, P. (1987): *Un método para la investigación-acción participativa*. Madrid: Editorial Popular.
- MARTÍN, D., PADILLA, T. y REBOLLO, A. (1991): «El modelo de investigación-acción como propuesta de análisis del proceso de formación-investigación de una escuela de padres». *Comunicación presentada al II Congreso de Investigación Educativa de Estudiantes*. Sevilla.
- MARTÍN, D., PADILLA, T. y REBOLLO, A. (1992): «La investigación-acción: una propuesta metodológica para la formación en una escuela de padres de Sevilla». *Comunicación presentada al II Simposio Internacional de Teoría Crítica e Investigación-Acción*. Valladolid.
- KEMMIS, S. y McTAGGART, R. (1988): *Cómo planificar la investigación-acción*. Barcelona: Laertes.

DIMENSIONES IMPLICADAS EN LA INSERCIÓN LABORAL: ANÁLISIS A PARTIR DE UNA ESCALA

Mercè Jariot García
Màrius Martínez i Muñoz
Departamento de Pedagogía Aplicada
Universidad Autónoma de Barcelona

INTRODUCCIÓN

El Área de M.I.D.E. del Departamento de Pedagogía Aplicada (Universidad Autónoma de Barcelona) lleva trabajando desde hace varios años (como una de líneas de investigación de su equipo) en los procesos de inserción laboral de personas en paro. Dentro de esta línea se elaboró una escala de inserción laboral (véase JARIOT, MARTÍNEZ, RODRÍGUEZ: 1991) a partir de la cual se han realizado ya estudios sobre la eficacia de programas de inserción (MONTANE: 1991), así como otros trabajos descriptivos relacionados con la temática citada.

A partir del marco teórico de la inserción laboral se han contemplado tres aspectos presentes en los procesos de inserción a la vida activa: los conocimientos, las destrezas y las actitudes de inserción laboral, en relación a la meta última que es la obtención de un puesto de trabajo.

En estudios anteriores se han comparado los resultados, obtenidos a partir de la pasación de la citada escala de inserción, entre colectivos que han seguido programas de información para la inserción y colectivos-control que no han recibido ayudas externas. En los mismos se constataron diferencias significativas entre ambos grupos atribuibles a los citados programas (véase MONTANE et al. 1992).

En este trabajo analizamos los diversos factores presentes en la escala a partir del análisis de regresión múltiple, con objeto de detectar aquellas dimensiones (V.I.) implicadas en el proceso de inserción, que mejor explican la obtención o no obtención de un trabajo (V.D.), así como otros factores relevantes para la inserción.

En concreto pretendemos dar respuesta a las siguientes cuestiones:

— ¿Qué dimensiones (ítems del cuestionario) explican la varianza de la variable criterio «haber encontrado empleo»?

— ¿Qué dimensiones explican realmente el hecho de que las personas tengan un alto nivel de autoestima profesional?

Para tal fin se ha utilizado el paquete de programas estadísticos para ciencias sociales (SPSS) en su versión para PC. La prueba utilizada ha sido un análisis de la regresión múltiple. Cuyos resultados presentamos a continuación.

ANÁLISIS Y DISCUSIÓN DE LOS RESULTADOS

Dos son, a nuestro parecer los criterios fundamentales en nuestro cuestionario, que deben ser analizados prioritariamente: estar trabajando, puesto que la investigación se enmarca en los procesos de inserción laboral y programas de orientación e información para los mismos, y también, como criterio anticipatorio, el confiar en encontrar empleo, es decir, todos aquellos aspectos de autoestima profesional que pueden mejorar dichos procesos.

En primer lugar hemos escogido como criterio la variable «estar trabajando», para intentar encontrar cuales son los predictores de encontrar un empleo. Mediante el cálculo de la regresión múltiple, que nos facilita la correlación (r), el coeficiente de regresión estandarizado (β), el coeficiente de contribución (C) y su porcentaje ($C\%$), obtuvimos los siguientes resultados, predictores de el hecho de emplearse, así como su contribución a dicho criterio:

— En primer lugar la variable «saber comportarse en una entrevista» es la que mejor contribuye a explicar el criterio (10.67%).

— En segundo lugar, también explica el criterio el «tener confianza en encontrar empleo» (7.3%).

— En tercer lugar, «estar dispuesto a aprender cosas nuevas de la profesión» resulta predictor de un 7.25% del criterio.

— En cuarto lugar, «aceptar otros trabajos» y «utilizar los servicios del INEM» (6.36%) son variables que también contribuyen a explicar el criterio «trabaja».

Además de estos cuatro predictores de encontrar empleo también destacan sobre el resto los siguientes: tener entrevistas de selección y/o pruebas pendientes (3.81%), tener elaborada una estrategia para buscar empleo (3.46%), sentir que los compañeros te vean como un buen profesional (3.23%), enviar curriculums (2.43%), pensar que puedes ofrecer más que los demás trabajadores (1.46%), conocer los diferentes tipos de contratos (1.25%) y conocer las ofertas de empleo (1.15%).

El resto de variables contribuyen en inferior medida a explicar el criterio encontrar empleo (vease cuadro nº 1).

A la luz de los resultados aportados en nuestro estudio podemos comprobar como la autoestima profesional, las actitudes de inserción, la flexibilidad y una actitud abiertamente activa frente a las novedades explican en buena parte la variabilidad de «encontrar un trabajo». Es especialmente significativo comprobar como dichas características son tenidas como valiosas cuando nos planteamos una intervención en orientación, desde planteamientos de «Educación para la Carrera», las actitudes de flexibilidad, y la capacidad para adaptarse a los constantes cambios (en este caso las novedades que puedan aparecer en una profesión) son valiosas características que se deben trabajar en los procesos de inserción. Por las limitaciones de este tipo de contribución nos vemos obligados a no extendernos más en este extremo.

Finalmente, podemos decir que, en cuanto a este criterio («trabajar»), hemos encontrado que el conjunto de ítems del cuestionario contribuyen en un 61.63% a explicarlo, queda pues por ver qué factores no presentes en el cuestionario se relacionan con el criterio mencionado.

Nos pareció importante, también, encontrar los predictores de la variable «confiar en encontrar empleo» que, como hemos visto, es un factor relevante respecto a «conseguir un empleo», y comprobar si podemos considerarla un criterio importante en los procesos de búsqueda de empleo; en este sentido los resultados obtenidos nos permiten afirmar que todo el conjunto de ítems contribuyen a su explicación en un 67.85%. Profundizando en la contribución de cada uno de los 28 ítems de la escala destacamos los 11 siguientes:

— No estar desanimado (11.55%).

— Aceptar horarios irregulares (10.60%)

— Conocer los distintos tipos de contrato (9.47%).

CUADRO N.º 1
CORRELACIÓN MÚLTIPLE DE LOS ÍTEMS DEL CUESTIONARIO CON EL FACTOR TRABAJAR

DIMENSIÓN: TRABAJAR				
ÍTEMS DEL CUESTIONARIO	r	beta	C	C%
Confianza en encontrar empleo	.224	.3261	.0730	7.30
Realizar curso de formación	-.001	-.1199	.0001	0.01
Práctica en psicotécnicos	-.139	.0573	.0080	0.80
Acepta otros trabajos	.198	.3214	.0636	6.36
Utiliza servicios INEM	-.019	.1225	.0636	6.36
Desplazamientos	-.084	-.2147	.0180	1.80
Currículums enviados	.098	.2479	.0243	2.43
Puede ofrecer más que los demás	-.107	-.1360	.0146	1.46
Empresas contactadas	.039	-.1169	.0046	0.46
Dedicado bastante a buscar empleo	.070	.0282	.0020	0.20
Acepta categoría inferior	.140	-.0228	.0032	0.32
Conocimiento calificación profesional	.000	.1271	.0000	0.00
Dispuesto a aprender cosas nuevas	-.256	-.2833	.0725	7.25
Desanimado por no encontrar empleo	.047	-.1031	.0048	0.48
Cree que es buen profesional	-.149	.0547	.0081	0.81
Entrevistas/pruebas pendientes	-.212	-.1799	.0381	3.81
Tener hecho el curriculum	-.041	.0538	.0022	0.22
Trabajo depende de interés o suerte	.034	.0994	.0034	0.34
Estrategia búsqueda de empleo	-.143	-.2418	.0346	3.46
Conocimiento tipos de contrato	-.070	-.1784	.0125	1.25
Horario irregular	.063	.0139	.0009	0.09
Dedicación búsqueda de empleo	.028	.1596	.0045	0.45
Comportarse en una entrevista	-.302	-.3535	.1067	10.67
Leer cosas relacionadas con profesión	-.049	.1684	.0083	0.83
Vías de búsqueda utilizadas	-.075	-.0124	.0009	0.09
Compañeros ven como buen profesional	-.178	-.1817	.0323	3.23
Conocimiento ofertas de trabajo	-.088	.1309	.0115	1.15
TOTAL			.6163	61.63

- Estar trabajando (6.41%)
- Estar siempre dispuesto a aprender cosas nuevas de la profesión (5.36%).
- Aceptar otros trabajos (3.34%).
- Dedicarse a buscar empleo (4.7%).
- Tener entrevistas y pruebas de selección pendientes (2.65%).
- Creer que uno es un buen profesional (1.56%).
- Conocer las ofertas de trabajo (1.38%).
- Leer cosas referentes a la profesión (1.06%).

El resto de ítems no se explicitan por ser escasa su contribución a la variable criterio (ver cuadro n.º 2).

CUADRO N.º 2
REGRESIÓN MÚLTIPLE DE TODOS LOS ÍTEMS DEL CUESTIONARIO RESPECTO A LA
CONFIANZA EN ENCONTRAR EMPLEO

FACTOR: confianza encontrar empleo.				
ÍTEMS DEL CUESTIONARIO	r	beta	C	C%
Trabaja	.224	.2863	.0641	6.41
Conocimiento calificación profesional	-.009	-.1083	.0010	0.10
Empresas contactadas	.050	.0079	.0004	0.04
Desplazamientos diarios	.051	.0857	.0043	0.43
Cree que es un buen profesional	.092	-.1693	.0156	1.56
Realizar curso de formación	.000	.1550	.0000	0
Desanimado por no encontrar empleo	.322	.3588	.1155	11.55
Dedicación búsqueda de empleo	.200	.2336	.0470	4.70
Utilización servicios INEM	.119	.0644	.0086	0.86
Aceptaría categoría profesional inferior	.102	.0786	.0080	0.80
Cree que puede ofrecer más que los demás	.045	.0499	.0022	0.22
Aceptaría otros trabajos	.141	-.2369	.0340	3.34
Tiene práctica en psicotécnicos	.014	-.0775	.0011	0.11
Currículums enviados	-.001	-.1830	.0002	0.02
Dispuesto aprender cosas nuevas prof.	.272	.1969	.0536	5.36
Aceptar horario irregular	.289	.3559	.1060	10.60
Entrevistas/pruebas pendientes	.108	.2457	.0265	2.65
Trabajo depende interés o suerte	.094	-.0369	.0035	0.35
Conocimiento tipo de contratos	.264	.3588	.0947	9.47
Estrategia búsqueda de empleo	-.050	.0577	.0029	0.29
Tener hecho el currículum	.044	-.1181	.0052	0.52
Dedicación búsqueda de empleo	.040	-.0491	.0020	0.20
Saber comportarse en una entrevista	.016	-.2817	.0045	0.45
Leer cosas de la profesión	.116	-.0914	.0106	1.06
Vías de búsqueda de empleo utilizadas	-.022	-.0796	.0018	0.18
Compañeros ven como buen profesional	.223	.2312	.0052	0.52
Conocimiento ofertas de empleo	.169	.0815	.0138	1.38
TOTAL			.6785	67.85

Volvemos a comparar como las actitudes de inserción, la flexibilidad y la apertura a las novedades, aspectos expresados en algunos de los ítems que más contribuyen a la explicación del criterio, son los factores más relevantes. En menor grado encontramos algunos de los conocimientos relativos a las ofertas o a aspectos de la profesión.

CONCLUSIONES

Queda por realizar un análisis de los 8 factores contemplados en el marco teórico de la inserción del cual partimos (Montané, 1991) que son: disponibilidad, dedicación, amplitud, madurez ocupacio-

nal, autoestima, atribución de causalidad, destrezas de inserción y conocimientos de inserción, con el propósito de averiguar cuáles son los ítems que más explicarán cada una de las dimensiones anteriormente citadas y si el nivel predictor de los mismos es elevado en cuanto al factor analizado.

Por último, debemos matizar que, dadas las limitaciones de esta comunicación sólo nos ha permitido aportar una mínima parte del estudio y comprobar como diversos factores son predictores de encontrar trabajo. Creemos que dichos factores son ciertamente importantes y que deben tenerse especialmente en cuenta en los procesos de inserción de trabajadores en paro, sobretudo en momentos como el actual en el que la inserción de dicho colectivo es uno de los objetivos prioritarios de todos los agentes implicados, entre ellos los investigadores y profesionales de la orientación.

BIBLIOGRAFÍA

- BISQUERRA, R. (1989): *Análisis multivariable* (Vol. I). PPU. Madrid.
- JARIOT, M., MARTÍNEZ, M. y RODRÍGUEZ, M. (1991): «*Elaboración y validación de una escala de inserción laboral*». Actas del Primer Congreso Internacional de Formación Ocupacional. U.A.B. Bellaterra.
- MARSOL, L. y MONTANE, J. (1991): «*La inserción laboral y el programa de información y búsqueda de empleo*» HERRAMIENTAS, 15, pp. 40 - 47. Madrid.
- MONTANE, J. (1991): «*Cómo tratar la inserción laboral*» HERRAMIENTAS, 14, pp. 36-41. Madrid.
- MONTANE, J. et. alt. (1992): *Los procesos de inserción y la influencia del programa PIBE*. Documento de uso interno. Universidad Autónoma de Barcelona.

EDUCACIÓN AMBIENTAL Y EVALUACIÓN EN EDUCACIÓN PRIMARIA

Juliana M. Correa Manfredi

Escuela Taller de Medio Ambiente de Castilleja de la Cuesta

M^ª Teresa Paniagua Aguilar

I.C.E. de Sevilla

Contemplar la Educación Ambiental en la Educación Primaria queda suficientemente justificado en la colección de materiales curriculares de la Consejería de Educación de la Junta de Andalucía para dicho nivel educativo, porque:

— La construcción de conceptos y actitudes por parte de las personas se desarrolla mediante un proceso que tiene sus inicios en las primeras percepciones y acciones que realiza el sujeto y que sirven de base para aprendizajes cada vez más diversos y profundos, estando éstos sujetos a contextos experienciales concretos.

— El proceso de construcción de conocimientos se da siempre en interacción con el medio ambiente, en todas sus facetas (sociales y culturales, biofísicas...) por lo que deberían de estudiarse en cada momento y circunstancia los contextos de aprendizaje de los que partir y trabajar con los alumnos según su edad, situación sociocultural, experiencias previas, intereses, etc.

Su introducción en el currículum se articula a través de la consideración de la misma no como una disciplina a impartir, sino como tema transversal que impregna al resto de las áreas del currículum, siendo las finalidades deseables de la misma comprender la multiplicidad del medio, resultante de las interacciones entre los diferentes sistemas que lo configuran (físico-químicos, biológicos y sociales), así como desarrollar las capacidades que posibiliten intervenir en él de forma positiva, solidaria y cooperativa.

Pero dirigirse hacia estas metas desde las edades más tempranas supone construir gradualmente la noción de medio, y por tanto delimitar la percepción que del mismo poseen los alumnos de este nivel, la cual diferenciaremos en cuatro enunciados, no entendidos éstos como etapas rígidas que preceden a la siguiente, sino como orientativas y flexibles para explicar las relaciones entre el niño y el medio. En un primer momento el alumno percibe el medio como una **realidad indiferenciada**, en la que objetos, sensaciones y relaciones aparecen unidos sin una organización definida, diferenciando y reconociendo progresivamente personas, objetos, y situaciones por su importancia para el niño (los padres, los amigos, el parque), llegando a establecer categorías muy básicas, como vivo-no vivo, bueno-malo. A medida que la experiencia del sujeto aumenta, el medio se va convirtiendo en un **escenario** (lugar donde ocurren cosas), ligado siempre a las vivencias del niño, que posee un carácter yuxtapuesto, en el que los objetos son percibidos de forma aislada y fragmentada (en el campo hay pájaros, árboles, flores...) y escasa o nulamente relacionados. Tiene por tanto esta concepción un marcado carácter

egocéntrico avanzándose, a medida que aparece la «necesidad por la sociedad», hacia una mayor descentración, considerando el medio como un conjunto de recursos relacionados entre sí y utilizados para mejorar las condiciones de vida de la colectividad; sitúa a los hombres en el centro de lo existente, pero a la vez en el exterior, ya que considera «lo humano» como algo distinto al «medio natural», no obstante comienza a establecer relaciones funcionales entre los distintos elementos que integran el medio. Se va avanzando por tanto hacia una visión compleja en la que se reconocen los diferentes componentes y se establecen las influencias de unos sobre otros, es decir, se concibe el medio como un **conjunto de sistemas que se relacionan** e interactúan.

Una vez realizadas estas consideraciones previas, decir que nuestra actuación se realizó en el primer curso de educación primaria (6-7 años), y que por tanto la concepción de medio a la que puede llegar el niño se encuentra limitada a su consideración como realidad indiferenciada y escenario. Se diseñaron actividades que contemplaban la dimensión ambiental para todo el curso, que aparecían con ese carácter de transversalidad en las áreas de conocimiento del medio, lenguaje, matemáticas y educación física, lo que suponía considerar la Educación Ambiental como principio didáctico que implicaba una serie de rasgos metodológicos a la hora de su puesta en práctica y que se traducía en crear situaciones de aprendizajes motivadoras y significativas para el niño, en la que éste participaba de forma activa, posibilitando el juego y la interacción con el medio.

No obstante, para concretar, nos centraremos en una sola de las actividades realizadas, considerada ésta como actividad tipo, ya que implicaba la realización de varias para la consecución de los objetivos, mostrando a través de la misma la evaluación realizada.

Nombre de la actividad tipo: «Recuperación del Árbol de Navidad».

Áreas con las que se relacionó:

1.— Educación Física. Técnicas y estrategias contempladas, relacionadas con el desarrollo psicomotor del alumno: distintos tipos de desplazamiento referidos a espacios, acampadas, manejo de instrumentos de la vida cotidiana, salidas y desplazamientos en diferentes momentos y circunstancias (ciclos estacionales, tormentas, construcciones nuevas,...), incidencia del hombre en el entorno, conocer y respetar a nuestros compañeros, riqueza en el trabajo de equipo (relación, comunicación, cooperación,...), posibilidades y limitaciones físicas, identidad socionatural en el propio medio (costumbres y tradiciones), utilización de todos los sentidos para percibir el medio, valoración y respeto del propio cuerpo en el medio, prevención de accidentes en la naturaleza, higiene corporal, cuidado y aseo (utilización de prendas adecuadas para cada tipo de actividad), relaciones con el medio natural y urbano...

2.— Matemáticas. Nociones y cuestiones relacionadas con la Educación Ambiental: el número como modo de acercamiento a la realidad, relaciones de igualdad y desigualdad, los agrupamientos de elementos, elaboración, conocimiento y utilización de estrategias de medida, reconocimiento e identificación de magnitudes (longitud, peso, etc.), desarrollo de sistemas de referencia (localización de objetos en el espacio), las formas en el espacio.

3.— Lenguaje. Considerando los contenidos relacionados con la Educación Ambiental en tres apartados: búsqueda de información (observación directa, norma de uso de la biblioteca, conseguir diversidad de información), organización de la información y expresión de la información (utilización de diferentes formas de representación lingüística).

4.— Conocimiento del medio. Se consideraron: respeto y conservación responsable del medio físico, los animales y las plantas, profesiones que repercuten en la conservación del medio, usos y abusos del hombre sobre el medio, reconocimiento de las estaciones del año, distinción de los

diferentes tipos de suelo (tierra), salud y medio ambiente (limpieza personal, prevención de enfermedades, suciedad en el ambiente (contaminación), efectos sobre la salud personal, actividades saludables (para el niño y para el medio).

Breve descripción de la actividad:

Aprovechando la proximidad de las fiestas de Navidad se preparó un guiñol, como elemento motivador, en el que se presentaban varios personajes (un ángel, un olivo, una palmera, un abeto, una pareja de pastorcitos), siendo la narradora, ataviada con traje de pastora, el único personaje real. Durante la representación cada árbol exponía que frutos daba y para qué servía, finalizándose la obra con la colocación, por parte del ángel, de estrellas sobre el árbol de Navidad, insistiéndose en que una vez finalizada la festividad no se tirasen los árboles al contenedor y entregando a los niños una tarjeta, en forma de abeto, en la que se ofrecía un teléfono y una dirección a la que se podía recurrir para que les recogiesen el árbol en lugar de tirarlo. Partiendo de este guiñol, y buscando como **objetivo final** el desarrollo y promoción de la salud individual y colectiva para que el niño sea capaz de intervenir en la construcción de su medio inmediato, así como reconocer y valorar su medio ambiente cercano como forma de integración social, fomentando el respeto por los elementos que lo rodean, se diseñaron actividades para cada una de las áreas y aspectos anteriormente señalados, intentando en todo momento globalizar la mayor parte de los mismos en cada una.

Evaluación:

Consideramos la actividad mencionada como una más de las integrantes de un programa de Educación Ambiental desarrollado a lo largo de todo el curso, por lo que a la hora de hablar de evaluación nos referimos de forma general a todo el programa para el desarrollo de presupuestos teóricos, si bien a lo largo de este apartado analizaremos cómo se evaluó la actividad tipo a la que nos hemos estado refiriendo a lo largo de esta exposición.

Tan importante en un programa de educación ambiental como la delimitación de objetivos, selección de los contenidos y planificación y diseño de las actividades, es la evaluación del mismo, concebida ésta como el proceso de recogida de toda la información necesaria para la interpretación y valoración sistemática de todos y cada uno de los elementos que intervienen en el desarrollo de dicho programa.

Si bien entendemos que la evaluación no sólo implica la valoración de los alumnos, y que ésta a su vez no supone identificar evaluación del alumno con evaluación del producto, ni equiparar resultados académicos con desarrollo cognoscitivo, si será en este aspecto de la evaluación en el que nos centraremos en la exposición de nuestra actividad; concluiremos por tanto que «no puede reducirse la evaluación educativa a evaluar al alumno; ni consiste la evaluación del alumno en evaluar sólo su rendimiento académico» (Abarca, 1989, p. 181), aunque es objetivo específico del profesor realizar dicha evaluación.

En el desarrollo de programas de educación ambiental está más generalizada la evaluación del proceso, tanto del diseño como de la ejecución, quizás porque estos datos sean más accesibles, sobre todo si tenemos en cuenta que hasta ahora los programas de educación ambiental realizados en los centros escolares partían de agentes externos al propio centro y su contacto con el mismo se limitaba al tiempo de implementación del programa. Sin embargo nos parecen igualmente importantes los resultados obtenidos por los alumnos, no sólo los referentes al ámbito cognitivo sino también los de ámbito procedimental y actitudinal, ya que son precisamente estos últimos los que en mayor medida se pretenden desde la educación ambiental. Dada la dificultad de evaluar estos aspectos en función de la

diversidad de variables posibles a analizar, es por lo que creemos conveniente el establecer un modelo de evaluación del alumno que abarque el mayor número posible.

A continuación describiremos la evaluación realizada en nuestra actividad «Recuperación del árbol de Navidad». Como fase previa a la puesta en práctica de la actividad pasamos a una lista de control para ver qué conductas de las que se pretendían desarrollar con la actividad, manifestaba el niño. Una vez analizados los resultados (evaluación inicial) usamos esas conductas como punto de partida para seguir un aprendizaje significativo, empleando un elemento motivador (el guiñol) acorde con sus características. Para evaluar la actividad nos basamos en la observación sistemática tanto directa como indirecta, haciendo uso de una escala de estimación y un registro anecdótico. Los aspectos observados fueron los siguientes:

— Reconocer diferentes tipos de árboles. Durante la representación del guiñol los diferentes árboles que intervenían iban exponiendo sus características, guiando la observación del niño, que era reforzada y mediada por el narrador.

— Identificación de la estación del año por las fiestas que se celebran. Se montó un árbol (artificial) de Navidad en la clase participando todos los niños, por grupos en las labores de decoración.

— Diferenciar los distintos tipos de suelo (tierra). Construyeron un terrario para distinguir los distintos tipos de tierra, comparando además la arena de playa y chinos utilizados en la maceta del árbol de la clase con la tierra usada para sembrar unas legumbres.

— Desarrollo de su vocabulario básico en términos ambientales. Este aspecto se estimó por medio de la invención de un cuento del que se le daba el punto de partida y un diálogo en el rincón de encuentro de los hechos sucedidos en el cuento.

A través del desarrollo de actividades en grupo se valoró:

— Conocimiento de la importancia que pueden tener los árboles para mantener el aire limpio.

— Importancia de la luz y temperatura para los árboles.

— Imposición de la conducta que cree correcta al resto de sus compañeros.

Se convocó una reunión con los padres (observación indirecta) para conocer:

— Aplicación, en el contexto familiar, de lo aprendido durante el desarrollo de la actividad.

— Cuidados y respeto procesados a los árboles de Navidad (de familiares y amigos).

Del registro anecdótico se encargó la maestra, anotando en él todo lo que ella veía fuera de lo común en las conductas habituales de los niños, así como las diferencias de comportamientos manifestadas entre ellos. Si se han expuesto estos aspectos evaluados y no algunos más relacionados con otras áreas es porque éstos fueron los conjuntamente valorados por la profesora y los observadores externos.

Si comparamos el modelo de evaluación utilizado para todo el programa con el usado por otros programas de intervención, en un intento de encuadrarlos dentro de un modelo teórico, concluimos que estaría más próximo al modelo de evaluación crítico (Colas, 1992b), ya que el programa persigue la transformación de los destinatarios, suponiendo este cambio un paso para transformaciones más globales, que repercutirán a nivel personal y grupal, e incluso a más largo plazo, comunitario y social. En nuestro programa se evaluó tanto el proceso como el resultado, sin olvidar el contexto del que se partía, estando el evaluador (en nuestro caso evaluadores) implicados en todo el proceso y comprometidos tanto en el diseño y desarrollo, como en la obtención de resultados, ya que al trabajar con un grupo de tan corta edad (6-7 años) eran conscientes de lo que su ejemplo podía suponer para la consecución de los objetivos marcados. Por último señalar que la evaluación fue un proceso en sí, ya que se realizó de forma continua a lo largo de todo el programa produciendo cambios en el mismo en la medida que eran necesarios, con lo que poseía un carácter cíclico.

BIBLIOGRAFÍA

- ABARCA PONCE, M. P. (Coord.) (1989): *La evaluación de programas*. Madrid. Editorial Escuela Española.
- CABRERA, F. y ESPÍN, J. (1986): *Medición y evaluación educativa*. Barcelona. P.P.U.
- CARMEN, L. DEL y ZABALA, A. (1991): *Guía para la elaboración, seguimiento y valoración de proyectos curriculares de centro*. Madrid. C.I.D.E. Centro de Publicaciones, M.E.C.
- CARR, W. y KEMMIS, S. (1988): *Teoría crítica de la enseñanza*. Barcelona. Martínez Roca.
- COLÁS, P. (1992a): «La evaluación de programas desde la perspectiva crítica». Comunicación presentada al *Simposio Internacional sobre teoría crítica e investigación/acción*. Celebrado en Valladolid del 2 al 4 de marzo.
- COLÁS, P. (1992b): «Evaluación de programas: modelos y enfoques metodológicos». *Cuestiones pedagógicas*. Nº 8 y 9, pp. 411-423.
- GIMENO, J. y PÉREZ GÓMEZ, A. (1983): *La enseñanza: su teoría y su práctica*. Madrid. Akal Universitaria.
- JUNTA DE ANDALUCÍA (1992): *Colección de materiales curriculares para la educación primaria*. Sevilla. Consejería de Educación y Ciencias.
- ROMÁN PÉREZ, M. y DíEZ LÓPEZ, E. (1990): *Un modelo de diseño curricular de aula en el marco de la reforma*. Currículum y aprendizaje. Madrid. Unidad de Programas Educativos. M.E.C.
- SOBRADO FERNÁNDEZ, L. (1992): *Intervención Psicopedagógica y orientación educativa*. Barcelona. P.P.U.

EFFECTOS DIFERENCIALES DE PROGRAMAS DE EDUCACIÓN INFANTIL EN FUNCIÓN DE SU GRADO DE ADECUACIÓN A LOS PRINCIPIOS ORGANIZATIVO/DIDÁCTICOS ESTABLECIDOS EN EL DISEÑO CURRICULAR BASE

M^a Teresa Aguado Odina
Universidad Nacional de Educación a Distancia

1. INTRODUCCIÓN

La presente comunicación analiza los efectos diferenciales de programas de Educación Infantil en función de su grado de adecuación a los principios organizativo/didácticos establecidos en el Diseño Curricular Base (MEC, 1989) y forma parte de una investigación más amplia en la que se han analizado y valorado comparativamente diversos proyectos de Educación Infantil desarrollados en el marco del Plan Experimental para la Reforma de la Enseñanza y proyectos no pertenecientes a dicho Plan.

Una vez identificadas las variables de entrada y proceso que establecen diferencias entre los dos grupos de programas y que se relacionan con los resultados obtenidos por los alumnos, se analizan los efectos diferenciales de los programas en función de su pertenencia al Plan Experimental, grado de adecuación a los principios organizativo-didácticos, establecidos en el Diseño Curricular Base y modelo educativo ejemplificado, adoptando como criterio los objetivos generales fijados para la Reforma de la Educación Infantil (MEC, 1986/89).

2. DISEÑO CURRICULAR BASE (DCB)

La Reforma educativa tiene por objeto no sólo, ni principalmente, la ordenación o estructuración de la educación. Se refiere, sobre todo, a los contenidos, a la oferta y oportunidades de experiencias de aprendizaje que la escuela ofrece a los alumnos. Una reforma de la enseñanza es siempre reforma del currículo, es decir, de las intenciones educativas y de las condiciones para su efectividad. (MEC, 1989:7). El planteamiento curricular adoptado por el MEC es el de un Diseño Curricular Base (DCB) abierto y flexible, que ha de quedar ulteriormente determinado en sucesivos niveles de concreción por las Comunidades Autónomas con competencias educativas, los centros y los propios profesores.

El DCB representa el nivel más general del diseño curricular y tiene carácter obligatorio. En él se señalan la estructura, los objetivos generales, los principios organizativo-didácticos y la estructura curricular de la etapa 0-6 años. El DCB no se inclina por un método determinado por considerar que

no existe el buen método por definición. En su lugar establece una serie de principios que han de orientar la organización y desarrollo de los procesos de enseñanza-aprendizaje. Estos principios se ejemplifican por procesos en los que la conducta del profesor está orientada a lograr la adaptación a los alumnos, fomento de la actividad, experimentación y cooperación de los alumnos; los contenidos, globalizados y diversificados; los espacios y materiales disponibles suficientes, variados y organizados sistemáticamente; la temporalización flexible; la evaluación continua realizada con instrumentación adecuada; participación frecuente y activa de los padres y equipo docente integrado y activo.

3. PLAN EXPERIMENTAL

Entendiendo que los efectos diferenciales constatados al comparar los programas del grupo reforma y grupo no reforma pueden ser explicados en gran medida en función de los procesos reales didáctico-organizativos desarrollados en la clase, se plantea como **hipótesis** a comprobar.

— Existen diferencias estadísticamente significativas en los resultados obtenidos a lo largo de los dos cursos académicos correspondientes al segundo ciclo de educación infantil por los alumnos de programas caracterizados por su mayor grado de adecuación al Diseño Curricular Base (DCB) cuando se comparan con los de alumnos de programas menos adecuados al mismo si se adopta como criterio el grado de consecución de los objetivos generales establecidos para la Reforma de la Educación Infantil.

Se plantea un **diseño cuasi-experimental**, que es el que más garantías de control ofrece en los contextos educativos reales y por considerar que es el que mayor susceptibilidad de aplicación tiene en el presente estudio. Como en todo diseño cuasi-experimental se va a controlar el cuando y el a quién de la medición, aunque no la programación de estímulos experimentales. Se trata de un diseño jerárquico de seis grupos intratratamientos, transversal —se han tomado medidas en momentos dados del proceso— y longitudinal de tratamiento —se realizan medidas en seis momentos diferentes a lo largo de los dos cursos correspondientes al segundo ciclo de escuela infantil— y análisis de variables tanto estructurales como procesales.

Las **variables** de entrada son personales —referidas a padres, profesores y alumnos— y contextuales —estructurales y dinámicas—. Las variables de proceso en relación con el grado de adecuación al DCB describen la conducta del profesor, contenidos, espacios, materiales, temporalización y evaluación. Las variables criterio se agrupan en las tres grandes áreas consideradas en los objetivos generales para la Reforma: cognoscitivo-verbal, afectiva y psicomotriz.

La **muestra** ha quedado constituida por dos grupos de programas en función de su grado de adecuación al DCB medido mediante la aplicación de la «Escala para la observación de procesos organizativo-didácticos». Un grupo (G+) ha quedado constituido por los programas —grupos o clases— cuya puntuación en la escala es superior al 50% del total máximo posible. El segundo grupo (G-) lo constituyen los programas con puntuaciones inferiores a este 50%. Así, la muestra final se distribuye como sigue: 66 alumnos (35,9%) pertenecientes a cuatro centros —dos integrados en el Plan Experimental para la Reforma de la Educación Infantil y dos ajenos al mismo— en el G+ y 118 (64,1%) pertenecientes a dos centros del Plan Experimental en el G-.

4. INSTRUMENTACIÓN: «ESCALA PARA LA IDENTIFICACIÓN DEL GRADO DE ADECUACIÓN DE LOS PROCESOS DE CLASE A LOS PRINCIPIOS ORGANIZATIVO/DIDÁCTICOS DEL DISEÑO CURRICULAR BASE (DCB)»

La escala elaborada tiene por objeto medir la variable «adecuación al DCB» y permitir el registro sistemático de los procesos didáctico-organizativos desarrollados en la clase durante amplios períodos

de tiempo a lo largo del curso. La escala consta de ciento dos ítems seleccionados y organizados en función de las orientaciones contenidas en el Diseño Curricular Base (MEC, 1989) y de aportaciones procedentes de trabajos inscritos en enfoques teóricos próximos a dichos planteamientos inscritos en enfoques teóricos próximos a dichos planteamientos (Saunders y Bingham-Newman, 1987). El proceso seguido en la elaboración del instrumento ha sido el siguiente:

1. Elaboración de una lista de 120 ítems correspondientes a todas las dimensiones relevantes en función de los presupuestos teóricos —conducta del profesor, contenidos, espacios, materiales, horario, evaluación, equipo docente, participación padres— y acordes con los principios y estrategias didácticas asumidas —globalización, fomento de la autonomía y la actividad, significatividad del aprendizaje, participación del alumno, adaptación a los intereses y ritmos del grupo, horarios flexibles, materiales y espacios variados y organizados, trabajo en equipo de los profesores, evaluación continua y participación de los padres—.

2. Presentación de los ítems a diez jueces vinculados a la educación preescolar o/y básica e investigadores educativos.

3. En función de los resultados de este examen, se eliminaron o modificaron algunos ítems, quedando finalmente la escala constituida por los 102 ítems de la versión definitiva.

Dicha versión final está compuesta de ocho subescalas, cada una de las cuales consta de ítems que corresponden a acciones y procesos fácilmente observables referidos a conductas del profesor, contenidos, espacios, materiales, horario, evaluación, relaciones con los padres y equipo educativo. El nivel de inferencia de los ítems es bajo para facilitar la observación y aumentar la validez de los registros. La fiabilidad del instrumento se ha determinado en función del grado de acuerdo entre observadores. Las puntuaciones de porcentaje de acuerdo entre jueces es de 8 para la escala total. Los porcentajes de acuerdo para cada una de las dimensiones de la escala oscilan entre el 1 correspondiente a relaciones con los padres y el 7 de contenidos, horario y equipo docente. Otros porcentajes son 9 de materiales, el 8 de conducta del profesor, espacios y evaluación.

5. RESULTADOS

Los resultados obtenidos confirman la existencia de diferencias significativas en algunas de las variables criterio analizadas cuando se comparan los resultados de programas agrupados en función del grado de adecuación de los procesos observados en clase a los principios organizativo-didácticos establecidos en el DCB. Los datos obtenidos nos permiten concluir que:

a) En el **área cognoscitivo-verbal**, los programas que presentan una mayor adecuación a los principios organizativo/didácticos establecidos en el DCB son más eficaces que los menos adecuados en la obtención de superiores resultados al final de los dos cursos estudiados en complejidad y diversidad de pensamiento, capacidad para las matemáticas y la música y creatividad medida mediante observación directa en la familia. Los programas con un menor grado de adecuación al DCB, por su parte, obtienen superiores resultados en capacidades básicas medidas mediante pruebas altamente estructuradas —capacidad intelectual general, agudeza perceptiva, capacidad numérica y memoria—. Sus efectos positivos se manifiestan ya al final del primer curso y se mantienen a lo largo del segundo curso estudiado.

b) En las variables del **área afectiva**, los programas que ejemplifican procesos más adecuados a los principios organizativo/didácticos establecidos en el DCB son más eficaces en el desarrollo de actitudes y hábitos de independencia, iniciativa, autonomía y concentración. Frente a esto, son los programas menos ajustados al DCB los que obtienen superiores resultados en el desarrollo de habilidades sociales medidas mediante observación directa en la escuela. No se constatan efectos diferenciales en autoconcepto, actitud hacia la escuela, habilidades sociales en la familia y hábitos de persistencia y autocontrol.

c) Por lo que se refiere a las variables del **área psicomotriz**, no se constatan efectos diferenciales en función del grado de adecuación al DCB de los procesos observados en clase cuando se utiliza una escala de observación directa en la familia, pero sí —a favor del grupo menos adecuado al DCB— cuando se utiliza una prueba estandarizada altamente estructurada.

El análisis de estos resultados debe tener en cuenta las limitaciones del instrumento utilizado para identificar el grado de adecuación de los procesos de clase a los principios organizativo-didácticos del DCB, el cual incluye dimensiones y conductas significativas, pero no de forma exhaustiva. Sería de interés establecer comparaciones entre programas con puntuaciones altamente divergentes en la escala, ya sea considerada globalmente como atendiendo a dimensiones específicas.

REFERENCIAS BIBLIOGRÁFICAS

- AGUADO ODINA, M. T. (1991): *Análisis y valoración de programas pertenecientes al Plan Experimental para la Reforma de la Educación Infantil desarrollados en Zaragoza-capital*. UNED. Madrid.
- ALKIN, M. C. (1990): «Curriculum evaluation models», en H. J. WALBERG y G. D. HAERTEL (eds.): *The International Encyclopedia of Educational Evaluation*. Pergamon Press. Oxford. 166-168.
- CARBALLO, R.; GARCÍA, M. y OLIVEROS, C. (1988): «Educación infantil: relación entre estilo de enseñanza y estilo cognitivo». *Actas del IX Congreso Nacional de Pedagogía*. Alicante, 27 septiembre a 10 octubre, 318-319.
- CARMENA, G. y otros (1989): *Niveles de desarrollo de la población al acceder al Ciclo Inicial*. CIDE. Madrid.
- CLARK, M. y CHEYNE, W. M. (eds.) (1979): *Studies in pre-school education* Hodder and Stoughton. London.
- DECKER, C. A. y DECKER, J. R. (1984): *Planning and Administering Early Childhood Education*. Merrill Publishing Company. Columbus. Ohio.
- DE MIGUEL, M. (1988): *Preescolarización y rendimiento académico*. CIDE. Madrid.
- FRANKLIN, M. B. y BIBER, B. (1977): «Perspectives and early childhood education: some relations between theory and practice», en L. G. Katz (ed.): *Current topics in early childhood education*. Vol. 1. Ablex-Norwood. New Jersey.
- GOODWIN, W. L. y DRISCOLL, L. A. (1984): *Measurement and Evaluation in Early Childhood Education*. Jossey-Bass. San Francisco.
- KAMII, C. (1985): «Leading Primary Education Toward Excellence», en *Young Children*. Vol. 40, nº 6, pp. 3-9.
- KAMII, C. y DE VRIES, R. (1974a): *Piaget-based curricula for early education: three different approaches*. Society for Research in Child Development. Philadelphia.
- MAYER, R. S. (1971): «A Comparative Analysis of Preschool Curriculum Models», en *As the Twig is Bent*. Houghton Mifflin. Boston, pp. 286-314.
- M.E.C. (1985): *Documento de trabajo - I*. (Programa Experimental de Educación Infantil). Madrid.
- M.E.C. (1986): «Hacia un nuevo modelo de educación infantil». *Actas de las Jornadas de análisis y revisión de los Programas Renovados de Preescolar*. Madrid.
- M.E.C. (1986): *Anteproyecto de Marco Curricular para la Educación Infantil*. Madrid.
- M.E.C. (1987): *Proyecto para la Reforma de la Enseñanza*. Madrid.
- M.E.C. (1989): *Diseño Curricular Base. Educación Infantil*. Madrid.
- M.E.C. (1992): *Nuevas orientaciones curriculares*. Madrid.

- ORDEN, A. de la (1986): "El currículum en la educación preescolar", en SÁNCHEZ CERREZO (dir.): *Enciclopedia de la Educación Preescolar*. Tomo 1. Diagonal/Santillana. Madrid, 85-96.
- OSBORN, A. F. y MILBANK, J. E. (1987): *The effects of Early Education*. Clarendon Press. Oxford.
- PETERS, D. L.; NEISWORTH, J. T.; YAWKEY, T. D. (1985): *Early Childhood Education: From Theory to Practice*. Brooks/Cole Publishers Company. Monterey. California.
- ROOPNARINE, J. L. y JOHNSON, J. E. (1989): *Educational models for Young Children*. Merrill C. Columbus.
- SAUNDERS, R. y BINHAM-NEWMAN, A. (1987): *Proyecto 0-6 de Educación Infantil. Informe piagetiano*. MEC. Madrid.
- STEVENS, J. H. y KING, E. W. (1987): *Administración de programas de educación temprana y preescolar*. Trillas. México.

INTERVENCIÓN ESTRUCTURAL EN UN CASO DE RETRASO ESCOLAR Y DÉFICIT EN COMPORTAMIENTO ADAPTATIVO

Amparo Martínez Sánchez
Universidad de Valencia

1. EL ESTUDIO DE CASOS EN LA INVESTIGACIÓN SOBRE LA INTERVENCIÓN EDUCATIVA DIFERENCIAL

El estudio de casos es uno de los modelos de investigación utilizados en las Ciencias Sociales y ha sido tradicionalmente el enfoque más utilizado en la investigación clínica. Su diseño es particularmente apropiado para el estudio de situaciones que requieren un notable grado de intensidad y, normalmente, un reducido período de intervención.

Los autores que tratan de este método de investigación se refieren comunmente al «análisis holístico e intensivo de un suceso particular que se produce en un determinado contexto y que requiere múltiples fuentes de información». (Guba y Lincoln, 1981; Bogdan y Bilken, 1982; Walker, 1983; MacClintock y col., 1983; Adelman y col., 1984; Ying, 1984; Merrian, 1988).

Tomando como punto de referencia las aportaciones de los principales autores que tratan el tema, señalamos siete notas como características del método del «estudio de casos»: 1) Remite a una referencia, a una unidad de estudio individual, el caso es un ejemplo de algo. 2) Exigencia de un examen holístico, intensivo y sistemático. 3) Necesidad de obtener información desde múltiples perspectivas. 4) Necesidad de considerar las características del contexto. 5) Tiene un carácter activo y dinámico. 6) Permite identificar los distintos procesos interactivos que caracterizan la situación estudiada. 7) Supone una estrategia encaminada a la toma de decisiones.

En resumen, podemos decir que el estudio de casos es una forma privilegiada en la investigación ideográfica y, particularmente, en la que tiene como finalidad la intervención educativa diferencial. Afronta el estudio de la realidad mediante un análisis detallado de sus elementos y de la interacción que se produce entre ellos y su contexto para llegar, mediante un proceso de síntesis, a la búsqueda del significado y a la toma de decisión que se requiere en la situación estudiada.

Este es el modelo de investigación elegido en nuestro trabajo. El objetivo que nos proponemos se cifra en la presentación del caso objeto de estudio y en la descripción del proceso seguido en el Programa de intervención estructural empleado en la situación concreta.

El caso práctico de Intervención Educativa que se presenta está dentro del ámbito de los Servicios Sociales de la Comunidad Valenciana, se ha desarrollado a partir de noviembre de 1992 y responde a una situación real en la que se enmascaran los datos de identificación por razones de respeto a las personas. En el momento de terminar esta Comunicación el caso no se ha dado por cerrado, por lo que

se ofrece el estado actual de la intervención y un avance de los principales resultados obtenidos hasta el momento.

2. LOS PROGRAMAS DE INTERVENCIÓN ESTRUCTURAL

Los programas de intervención estructural se enmarcan en la Teoría General de Sistemas propuesta por Bertalanffy (1950). De esta teoría se derivan una serie de principios entre los que básicamente destacamos cinco: 1º) Todo sistema constituye un todo organizado compuesto por elementos que son necesariamente interdependientes. 2º) Las influencias entre los elementos de un sistema son circulares más que lineales. 3º) Los sistemas tienen aspectos homeostáticos que mantienen la estabilidad de sus comportamientos. 4º) La evolución y el cambio se dan de forma inherente en los sistemas abiertos. 5º) Los sistemas complejos están compuestos por subsistemas.

La aplicación de este enfoque a las Ciencias Sociales conduce a la afirmación de que los fenómenos sociales deben ser considerados en términos de «sistemas», hecho que implica aceptar que existe una interrelación entre todos los elementos constituyentes de la sociedad y que los acontecimientos sociales parecen envolver algo más que las decisiones y acciones individuales y están determinados más bien por sistemas socioculturales: organizaciones, grupos de presión, prejuicios, ideologías, tendencias sociales, etc.

Desde esta perspectiva se pone el acento en la interpretación holística y sistémica de la conducta y la consideración de las variables internas y externas como factores determinantes del comportamiento. El interés de los programas de intervención se traslada del individuo al sistema relacional del que forma parte. Los factores esenciales de los problemas deben ser considerados como componentes interdependientes de un sistema total y, consiguientemente, el centro de atención pasa de los problemas y los comportamientos individuales a las interacciones que se producen entre los diversos elementos de los sistemas de los que forma parte el sujeto: la familia, la escuela, el barrio, la pandilla de amigos, etc. (Andolfi, 1988).

Se destaca así mismo la notable importancia que tiene la consideración de la ecología del desarrollo humano (Bronfenbrenner, 1987). El «ambiente ecológico» se concibe, topológicamente, como una disposición seriada de estructuras concéntricas, en la que cada una está contenida en la siguiente. Estas estructuras se denominan micro-, meso-, exo- y macrosistemas. Es decir, se consideran decisivos los patrones de actividades, los roles y relaciones interpersonales que la persona en desarrollo experimenta en su entorno más próximo que viene definido por características físicas y materiales determinadas; pero además tienen también interés las interrelaciones producidas en los entornos en los que la persona, de una u otra forma participa activamente y los contextos más amplios regidos por ideologías corrientes de pensamiento y tendencias que provocan situaciones, acontecimientos y hechos que afectan a lo que ocurre en el entorno que comprende a la persona en desarrollo.

Operativamente el estudio y tratamiento de los problemas que aparecen en el comportamiento de los sujetos ha de superar las limitaciones de los procedimientos analíticos y, consiguientemente, ha de centrarse en el conocimiento y la comprensión de las interacciones que se producen en los sistemas humanos de los que forma parte el individuo y en la determinación de las características específicas de los contextos en los que ocurren. Se hace necesario llegar a una comprensión significativa de las situaciones en las que el individuo se desenvuelve para orientar la intervención hacia el conjunto de estructuras que configuran el espacio vital en que el individuo se desenvuelve. A partir de estos planteamientos parece evidente que la familia es el contexto significativo más próximo al individuo que ayuda a explicar y, en su caso, a modificar los comportamientos desajustados del sujeto.

Este es el marco teórico y metodológico que en líneas generales preside el caso concreto que se presenta. No obstante, dadas las características y los déficits comportamentales que se aprecian en la

situación estudiada, entendemos que en aspectos muy concretos de la intervención, se hace preciso llevar a cabo una aproximación multimétodo que incluye algunas intervenciones puntuales, dirigidas a la adquisición de comportamientos adaptativos muy específicos, o a la supresión de conductas disruptivas, aspectos que coyunturalmente se consideran necesarios para posibilitar el proceso general que básicamente está inscrito en un programa sistémico dirigido a los elementos y estructuras más significativas del «ambiente ecológico» que configura las relaciones y la vida de los sujetos.

3. EL CASO DE «LOS MONTERO»

El caso de los niños que en adelante llamaremos Felipe y Adolfo, miembros de la familia Montero, tiene gran interés porque su problemática constituye un buen ejemplo de niños, pertenecientes a una familia disfuncional, con múltiples tensiones y maltratos, que presentan notables niveles de retraso y fracaso escolar, dificultades en la integración escolar y social y deficit y en los comportamientos adaptativos.

3.1. Descripción del caso

La familia Montero está formada por el padre, de 44 años, la madre de 38 y niños Felipe de 14 años y Adolfo de 8. La familia reside en una vivienda social situada en los alrededores de una gran ciudad, en un barrio de exclusión social que se caracteriza por la degradación de los espacios y la carencia de equipamientos de todo tipo. Los padres pertenecen a familias de emigrantes, situadas en los límites de la pobreza; el padre trabaja como peón en una fábrica, la madre además de atender a las labores de la casa realiza trabajos eventuales de servicio doméstico; ambos pueden leer y firmar pero son prácticamente analfabetos y presentan notables deficiencias de lenguaje. Los niños están escolarizados en el Centro de Enseñanza Estatal que funciona en el barrio.

Como consecuencia de una denuncia de malos tratos del marido hacia la mujer y el hijo mayor, la familia ha entrado en contacto con los Servicios Sociales que, teniendo en cuenta la dinámica relacional de la familia, las carencias educativas que tienen los niños y su falta de integración escolar y social, han solicitado ayuda técnica para mejorar la situación y, más específicamente, para superar las deficiencias y problemas que presentan Felipe y Adolfo.

3.2. El programa de intervención

La intervención se inicia con un primer contacto con los niños que vienen acompañados por la madre. En esta fase se observan y registran los primeros datos generales sobre los sujetos y se recaba información sobre los problemas que plantean en su vida familiar, escolar y social; también se obtiene una primera impresión sobre la configuración y estructura familiar y los problemas relacionales y comportamentales que a primera vista parece que afectan a padres e hijos. A este contacto inicial siguen una entrevista individual con cada niño y dos encuentros familiares en uno de los cuáles se consigue, después de múltiples intentos, que también esté el padre. La información obtenida a través de los encuentros reseñados, se completa con dos visitas a la casa familiar y con los contactos establecidos con los Profesores de los niños, el Psicólogo y el Director del Centro Escolar y los Trabajadores Sociales encargados de los Servicios Sociales en el barrio.

3.2.1. Fase de diagnóstico

A través de las diversas entrevistas, encuentros, observaciones y consulta de informes, notas y

otros datos escolares, se realiza un diagnóstico inicial tanto de la situación de la familia Montero, como de los niños, en el que destacan los siguientes aspectos:

En primer lugar hay que señalar la incompetencia parental que se pone de manifiesto en la incapacidad para conducir la vida cotidiana y resolver los conflictos familiares de forma breve, educativa y no violenta. El padre tiende a percibir la conducta de los niños, especialmente la de Felipe, como aversiva y provocadora y responde con malos tratos y con agresiones de todo tipo. La madre presenta un cierto grado de retraso mental que hace que atienda con poca organización y coherencia el trabajo doméstico —habitualmente no prepara una comida formal, se levanta tarde haciendo que niños falten al colegio o no sean puntuales, no dispone adecuadamente la ropa de los niños, realiza una mala distribución de tiempo/trabajo, es bastante desordenada, etc.—. Ante esta situación el padre responde ausentándose de casa en los ratos libres, bebiendo con alguna frecuencia y pegando y maltratando a la madre. Como consecuencia de estos comportamientos la mujer tiene miedo y presenta serios problemas emocionales de carácter depresivo que la conducen a un gran aislamiento social. Las relaciones de la familia con el vecindario no son buenas, frecuentemente tienen discusiones y peleas por asuntos triviales relacionados con la convivencia cotidiana.

Las actuaciones de los padres con los niños se caracterizan por la inconsistencia, la incertidumbre y la falta de concordancia entre las actuaciones de ambos; como consecuencia la vida de los niños se ve marcada por el miedo, la dificultad de adquirir competencias sociales y de establecer relaciones positivas y satisfactorias con los demás; todo ello provoca un gran desequilibrio en la estabilidad emocional y cognitiva de los niños que da lugar a la aparición de síntomas depresivos, sobre todo en Felipe, en el que se constata un alto sentido de culpabilidad, alteraciones en el sistema atribucional, sentimientos de indefensión, tristeza, disminución de la autoestima, etc. etc. (Polaino-Lorente, 1988). A todo ello se añade la escasez de recursos económicos, el mal estado y las deficiencias higiénicas de la vivienda familiar, la ausencia de recursos culturales y educativos de la familia, su vulnerabilidad ante las estimulaciones y provocaciones del medio y la falta de respuesta de los padres a las repetidas llamadas que se les hacen desde la escuela —nunca acuden a las reuniones programadas ni a las reiteradas citaciones de los Profesores tutores de los niños— Este hecho dificulta notablemente la propuesta de intervenciones coherentes y compartidas con las acciones que, según el personal educativo del Centro, deberían llevarse a cabo en la educación de Felipe y Adolfo.

En segundo lugar hemos de señalar las características más destacadas de los menores. Felipe es un niño que en su aspecto físico delata un notable déficit de crecimiento, siendo su talla y peso inferiores a lo esperado por su edad; tiene un aspecto desaliñado y con falta de higiene personal; su salud no es buena y a lo largo de su vida ha sufrido diversas enfermedades. Por lo que respecta a los aspectos escolares, ha repetido 1º y 5º y actualmente es repetidor de 6º curso; es un niño de inteligencia media-baja y presenta serias dificultades de lenguaje y notable retraso escolar; tiene un comportamiento hiperactivo, con una gran dispersión en las tareas que se le plantean, aparentemente se interesa por muchas cosas y comienza muchas actividades que deja sin concluir; raramente hace o termina las tareas que le encomiendan en la escuela para que realice en casa y en clase tiene conductas desintegradas y frecuentes peleas con sus compañeros a quienes siempre culpa de los incidentes en los que permanentemente está envuelto; siempre se está moviendo por la clase, y presta muy poca atención a las explicaciones y orientaciones de los Profesores. En general es un niño incómodo y, aunque se propone mejorar, de hecho consigue muy pequeños logros que además son muy pasajeros. En la vida familiar presenta comportamientos desajustados: es desobediente, bastante desordenado tanto en su presentación personal como en el uso de las cosas, con frecuencia contesta mal, pega a su hermano y tiene enfrentamientos y conductas provocadoras sobre todo con su madre; ante su padre calla por miedo ya que frecuentemente es objeto de malos tratos y palizas por parte de este. Su comportamiento extrafamiliar es extrovertido, inicia múltiples relaciones que son muy poco estables; los contactos que

establece, tanto en la escuela con los compañeros como en la calle son pasajeros y casi siempre acaban en peleas en las que en muchas ocasiones participan también los padres. Últimamente se ha incorporado a una pandilla del barrio que, sobre todo los fines de semana, se dedica a cometer diversas fechorías, entra en frecuentes peleas con las bandas rivales, etc. etc. Pensamos que Felipe se encuentra a un paso de la delincuencia.

Adolfo presenta, también serios retrasos tanto en el desarrollo físico como en el intelectual y psicomotor así como en su proceso de socialización. Es un niño que reclama permanentemente la atención de su madre y las personas que le rodean; en casa no realiza las tareas más elementales como vestirse, lavarse o comer sólo. En la escuela, lleva dos años de retraso, no ha aprendido a leer, su lenguaje es muy deficitario y a pesar de que tiene dos años más que sus compañeros no sigue el ritmo normal del aula. Su comportamiento es bastante desintegrado: descuida el material escolar, no cumple las normas establecidas, es incorrecto en el trato con los demás, nunca realiza trabajos en equipo y difícilmente acaba las pequeñas tareas que se le encomiendan, tiene una atención muy dispersa y no tiene ningún interés por el trabajo que le propone la maestra. En el recreo no juega con los demás, pelea y pega a sus compañeros más pequeños. En casa lo tratan como si fuera un niño de menor edad, de manera que se ha convertido en un pequeño tirano que reclama atenciones permanentemente.

3.2.2. Fase de terapia

A partir del diagnóstico de la situación se proponen los siguientes objetivos: 1º) Mejora del ambiente familiar que permita el cese de los episodios de abuso. 2º) Incremento de la comunicación entre los miembros de la familia dentro de unas relaciones positivas y gratificantes. 3º) Adquisición de las competencias que permitan a los padres ejercer un control educativo, no punitivo sobre los hijos. 4º) Reducción de las conductas desintegradas de Felipe y Adolfo y adquisición de los comportamientos adaptativos más elementales. 5º) Establecimiento de las condiciones que favorezcan el desarrollo de las capacidades básicas para mejorar los niveles de rendimiento en la escuela. 6º) Favorecer las relaciones personales de los niños con sus compañeros integrándolos en grupos alternativos a las pandillas predelictivas. 7º) Vinculación y coordinación educativa de la familia con la escuela, Servicios Sociales y otras Asociaciones radicadas en el barrio. 8º) Realización de las gestiones y acciones necesarias para la mejora de las condiciones de vida de la familia en cuanto a vivienda y tiempo de dedicación a los hijos.

El proceso terapéutico, orientado a la consecución de las metas señaladas, se centra en la elaboración de un plan de acción, coordinado por la persona que dirige la intervención, en el que participa el personal del Centro Escolar y de los Servicios Sociales y en el que se pretende implicar a la familia. Este proyecto educativo, se dirige al conjunto de variables que configuran el contexto en el que desarrolla la vida de los Montero, emplea controles y refuerzos periódicos y básicamente se centra en tres dimensiones:

1º) *Dimensión familiar.* Comprende una serie de técnicas dirigidas a la mejora del clima familiar que se basan en: a) Entrenamiento para lograr la competencia parental en la resolución de conflictos y en la evitación de los malos tratos del padre y agresiones de todo tipo. b) Repetición de experiencias positivas de comunicación entre los miembros de la familia en las que se favorece la descarga de tensiones acumuladas en un clima distendido y satisfactorio.

2º) *Dimensión de los comportamientos de los niños.* Comprende dos líneas de acción: a) Manipulación de las contingencias ambientales para conseguir la modificación y adquisición de determinadas conductas terminales básicas en los comportamientos sociales y en los escolares. b) Proporcionar los apoyos necesarios dentro y fuera de la escuela para la mejora del rendimiento escolar. En todo caso se suministran refuerzos permanentes y/o periódicos por parte de todas las personas que de una u otra

forma están en contacto con los niños.: coordinador de la intervención, Profesores, Trabajadores Sociales y, dentro de lo posible, padres.

3º) *Dimensión contextual*. Pretende actuar sobre el ambiente que rodea a la familia a fin de romper el aislamiento en el que se encuentra y proporcionarle unas relaciones más satisfactorias, hecho que repercutirá en la mejora de todo el clima familiar. Se actúa en dos direcciones: a) Integración de los padres en la Asociación de vecinos y establecimiento de contactos mensuales con la escuela. b) Participación de los niños en actividades y grupos alternativos a las pandillas del barrio: integración en el Centro de Día de los Servicios Sociales, en el grupo de los scouts que funciona en la Parroquia, salidas los fines de semana con Familias Acogedoras, etc.

Como reflexión conclusiva, a la hora de cerrar esta Comunicación, hay que decir que estamos en un proceso social que trasciende la mera consideración de las variables comportamentales y escolares de los niños, que requiere una intervención estructural y la implicación de múltiples instancias sociales. Este proceso exige un tratamiento prolongado, en el que los logros se producen muy paulatinamente y en medio de grandes dificultades. No obstante, en la situación estudiada se ha conseguido una sensible mejora en la estructuración y organización del hogar y en las relaciones familiares; también se ha constatado la adquisición por parte de los niños de comportamientos adaptativos básicos y la mejora de los rendimientos escolares. El caso no se ha dado por cerrado.

4. BIBLIOGRAFÍA

- ADELMAN, C., JENKINS, D. y KEMMIS, S. (1983): Rethinking Case Study: Notes from the second Cambridge Conference. En SIMONS, H. (Ed.): *Towards a Science of the Singular*, (pp. 45-61). Norwick: CARE.
- ALARIO BATALLER, S. (1990): *Estudio de casos en terapia del comportamiento*. Valencia: Promolibro.
- ANDOLFI, M. (1988): *Terapia familiar*. Buenos Aires: Paidós.
- ARY, D., JACOBS, L. C. y RAZAVIEHA, R. (1987): *Introducción a la investigación pedagógica*. México: Interamericana.
- BEDARD, M. G., DELL'ANIELLO, P. y DESBIENS, D. (1991): *La méthode des cas*. París. Gaëtan Morin.
- BERTALANFFY, L. (1950): An outline of General System Theory. *British Journal of Philosophy of Science*, 139-164.
- BOGDAN, R. y BIKLEN, S. (1982): *Qualitative Research for Education: An Introduction to Theory and Methods*. Boston: Allyn and Bacon.
- BROMLEY, D. B. (1986): *The case-study. Method in Psychology and Related Disciplines*. N. York: J. Wiley & Sons.
- BRONFENBRENNER, U. (1987): *The Ecology of Human Development*. Massachusetts: Harvard University Press.
- CEREZO, M. A. (1990): Programa de asistencia psicológica a familias con problemas de abuso infantil. *Revista de Seveis Socials*, 11-12, 687-691.
- CHRISTENSSSEN, C., HANSEN, A. y MOORE, J. (1987): *Teaching and the case method: Instructor's guide*. Boston: Harvard Business School.
- GUBA, E. G. y LINCOLN, Y. S. (1981): *Effective Evaluation, Improving the Usefulness of Evaluation. Results trough Responsive and Naturalistic Approaches*. San Francisco: Jossey-Bass.
- KEMMIS, S. (1980): The imagination of the case and the invention of the study. En SIMONS, H. (Ed.): *Towards a Science of the singular*. (93-142); Norwick: CARE.
- MACCLINTOCK, C., BRANNON, D. y MAYNARD-MOODY, S. (1983): Applying the logic of

- Sample Surveys to Qualitative Case Studies: The Case Cluster Method. En VAN MAANEN, J.: *Qualitative Methodology*. Beverly Hill: Sage Pub.
- MAHER, CH. y ZINS, J. (1989): *Intervención psicopedagógica en los Centros Educativos*. Madrid: Narcea.
- MERRIAM, S. (1988): *Case Study Research Education. A Qualitative Approach*. San Francisco: Jossey-Bass.
- POLAINO-LORENTE, A. (1988): *Las depresiones infantiles*. Madrid: Morata.
- SCHIEFELBEIN, E. (1974): *Teoría, Técnicas, Procesos y Casos en el planeamiento de la Educación*. Buenos Aires: El Ateneo.
- SIVELMAN, R. y WELTY, W. (1990): *Teaching with cases*. Pace University.
- TEDESCO, P. H. (1974): *Teaching with case studies*. Boston: MA. Public Information Center. Federal Reserve Bank of Boston.
- WALKER, R. (1982): The use of case studies in applied research and evaluation. En HARTNETT, A. *The social Sciences in Educational Studies*. London: Heinemann.
- WALKER, R. (1983): La realización de estudio de casos en educación. Ética, teoría y procedimientos. En DOCKRELL, W. y HAMILTON, D.: *Nuevas reflexiones sobre la investigación educativa*, (pp. 42-82). Madrid: Narcea.
- WILSON, S. (1989): *A case concerning content: using case studies to Teach subject matter*. East Lansing: National Center for research on Teacher Education.
- WORELL, J. (1983): *Tratamiento de las dificultades educativas. Estudio de casos*. Madrid: Anaya.
- YING, D. (1984): *Case Study Research. Design and Methods*. London: Sage.

EL SEXO Y LA CLASE SOCIAL COMO DETERMINANTES DE LOS INTERESES PROFESIONALES EN EL ÚLTIMO CURSO DE SECUNDARIA¹

José Luis Gaviria Soto
Universidad Complutense de Madrid

INTRODUCCIÓN

En 1984 el Centro de Investigación, Documentación y Evaluación Educativa (C.I.D.E.), organismo dependiente del Ministerio de Educación y Ciencia, recibió del mismo el encargo de llevar a cabo la evaluación externa de la reforma de la educación secundaria. Con tal objeto, el Servicio de Evaluación de dicho organismo procedió a la recogida de todos aquellos datos necesarios para tal fin. Una vez terminada dicha evaluación y publicados sus resultados (Álvaro, 1988), se consideró que el banco de datos resultante era susceptible de ser sometido a análisis secundarios. El Instituto de la Mujer acordó con el CIDE llevar a cabo una investigación que pusiera de manifiesto, en un análisis secundario, cómo son los intereses profesionales en función del género y la relación de esos intereses con la clase social de procedencia. La presente investigación es por tanto fruto de dicho acuerdo.

LA PROFESIÓN Y EL SEXO

En el pasado, los papeles de la mujer y el hombre en la sociedad estaban claramente delimitados. Junto con las conductas definidas como propias de cada sexo, también se definían los oficios o profesiones que estaban permitidas o asignadas a cada género. El mismo concepto de profesión era un concepto sesgado. En primer lugar, no existía para la mujer de clase media otra ocupación propia distinta de las tareas domésticas. Sólo las mujeres de las clases bajas podían desempeñar trabajos fuera del hogar. Y eso eran oficios. Porque la profesión era algo asociado a algunos hombres, mientras que las mujeres podían practicar oficios, pero nunca profesiones. Esto nos da una clara definición de roles no sólo en función del sexo sino también de la clase social.

Consecuencia de lo anterior es que la formación de hombres y mujeres era absolutamente dispar. La mujer de clase media escasamente tenía acceso a la escuela, y cuando la escolarización se prolongaba más allá de lo estrictamente básico, por ejemplo en jóvenes de clases altas, la formación recibida

¹ El presente artículo es parte del libro del mismo título que será publicado próximamente por el Instituto de la Mujer.

era la «adecuada» para mujeres, centrada en tareas del hogar o en capacidades que suponían un «adorno» social, como la música o alguna otra actividad artística. La universidad estaba descartada para las mujeres porque éstas por definición no ejercían profesiones.

Todavía hoy en España, aunque legalmente la mujer y el hombre están equiparados, hay profesiones que están consideradas como predominantemente femeninas o masculinas.

No vamos a hacer una descripción estadística de la evolución de la presencia de la mujer en los distintos ámbitos profesionales en la España del último medio siglo, pero es evidente que el número de mujeres desempeñando funciones profesionales ha aumentado. Lo mismo puede decirse de la presencia de la mujer en la universidad. Si tomamos cantidades absolutas, el incremento de este número en los últimos treinta años ha sido espectacular. Por ejemplo, la tasa de actividad de hombres y mujeres ha pasado de 82.51% y 23.85% en 1965 a 67.6% y 32.1% en 1987 respectivamente. Es decir, la diferencia en la tasa de actividad entre hombres y mujeres ha pasado de un 58.66% a un 35.5% (Instituto de la Mujer, 1987).

En el caso de los estudios universitarios, del curso 1980-1981 a 1984-1985 el alumnado universitario aumentó en un 26.5% en general, mientras que el número de mujeres aumentó en un 40%. Pero tanto en un caso como en otro es preciso acudir a la distribución de esas cantidades. Es verdad que en números absolutos la presencia de la mujer en el mundo del trabajo y en las carreras universitarias ha aumentado. Pero existiendo estereotipos muy fuertes que determinan en cierto grado las respectivas ocupaciones de hombres y de mujeres. Así, el número de mujeres ocupando puestos directivos y técnicos de alta cualificación es muy inferior al de los hombres. Por ejemplo, en el curso 1984-85 sólo el 13% de los matriculados en las escuelas técnicas superiores eran mujeres. Sólo el 1.7% de las mujeres se matricularon en esas carreras. E incluso puede matizarse todavía más. Entre las carreras técnicas, existen algunas que tienen una proporción muy pequeña de mujeres, como por ejemplo Ingeniería Aeronáutica o Naval, mientras que en otras escuelas, como Agrónomos, Montes, o Informática las diferencias no son tan grandes. Esa tendencia a la selección de carreras o profesiones «femeninas» es también muy notable en la formación profesional. Para evitar la discriminación efectiva en contra de la mujer no basta con que más mujeres acudan a la universidad. El campo de estudio que elijan es importante. El simple hecho de aumentar el número de mujeres asistiendo a la universidad no supone que como grupo estén mejorando su estatus social en la misma proporción. Si las mujeres no aumentan su presencia en las carreras que están nutriendo los puestos directivos en la sociedad, su estatus como grupo no variará.

IMPORTANCIA DE LA IGUALDAD EN LA REPRESENTACIÓN DE LOS SEXOS EN LAS PROFESIONES

El cambio científico tecnológico que se está produciendo en el sistema productivo del mundo occidental está exigiendo esfuerzos de adaptación a la fuerza de trabajo. De hecho, para que nuestra sociedad pueda alcanzar niveles de competitividad que nos eviten el desempeñar un papel secundario en el mercado internacional, es preciso contar con un mayor número de científicos y técnicos. La proporción de mujeres que eligen este tipo de carreras es mucho menor que la de hombres. Contamos por lo tanto entre las mujeres con una gran reserva de talentos que pueden incorporarse a estas profesiones.

Otra razón de índole distinta es la de la igualdad establecida por la declaración de los derechos humanos y nuestra Constitución. No basta con que legalmente hombres y mujeres tengan los mismos derechos. Es un hecho que los puestos de alta responsabilidad técnica y de gestión se nutren de personas procedentes de un número muy determinado de carreras en las que las mujeres están sub-representadas. Eso quiere decir que como grupo las mujeres tienen menor capacidad de influir en las

decisiones que afectan a su vida. Como individuos, la renuncia a aspirar a ciertas profesiones les limita a desempeñar papeles o profesiones consideradas como secundarias o subordinadas.

Y desde un punto de vista social, la ausencia de las mujeres del ámbito de la ciencia y la tecnología resta riqueza a esos ámbitos, teniendo que renunciar la sociedad a las aportaciones de los talentos de una gran parte del cincuenta por ciento de sus miembros.

OBJETIVOS DE LA INVESTIGACIÓN

Si los hombres y las mujeres acaban desempeñando profesiones distintas, eso quiere decir que existen algunos mecanismos que producen esa diferenciación. Los intentos de descubrir esos mecanismos y las formas posibles de abordar su estudio son muchos. En este trabajo se pretende estudiar en qué medida los intereses profesionales se ven afectados por el sexo y la clase social.

EL MECANISMO DE FORMACIÓN DE LOS INTERESES PROFESIONALES

Este estudio se centra en un grupo restringido de profesiones, aquellas que requieren al menos la terminación de los estudios secundarios. La razón no es otra que el que este trabajo analiza los datos correspondientes a un grupo de más de siete mil alumnos de último curso de formación profesional, bachillerato experimental y bachillerato ordinario. Por ese motivo, casi todas las elecciones consideradas están relacionadas con las carreras universitarias en sus distintos grados.

En las siguientes líneas vamos a tratar de formular un esbozo de lo que podría ser una teoría de los intereses profesionales, con el solo fin de que nos sirva de guía para la comprensión del significado de los datos que se analizan más adelante.

En primer lugar, debiéramos distinguir entre intereses profesionales y vocación profesional. Afortunadamente, hoy ya no se entiende la vocación como una especie de destino personal e irrenunciable, que, dada la tendencia a la idealización de la adolescencia, se presentaba a los jóvenes como un oculto objetivo a descubrir, algo ligado al afianzamiento de la propia identidad y a la afirmación de la personalidad.

En segundo lugar, podemos decir que los intereses profesionales son la conjunción de una información con una valoración personal de la misma. Tenemos por tanto dos elementos importantes. El primero de origen externo, la información. Está claro que no podemos sentir interés por aquello que desconocemos. No pueden existir en una persona intereses profesionales definidos, si no conoce en cierto grado el conjunto de las profesiones que se le ofrecen como posibilidades. El segundo elemento es de origen interno. La información recibida no se convierte en intereses profesionales, positivos o negativos, hasta que esa información es valorada. Este elemento interno supone también la existencia de una cierta madurez. No podemos hacer valoraciones si no poseemos capacidad de valorar, de dar valor. Por supuesto, esta capacidad está íntimamente relacionada con todas las experiencias previas de los individuos. Por lo tanto, los intereses profesionales, como tantas otras cosas, van formándose paulatinamente, y son uno de los productos de la educación. No sólo de la educación formal recibida en el entorno escolar, sino de todas las experiencias vitales del individuo, en el entorno familiar, en las relaciones con otros compañeros, de los valores e imágenes ideales transmitidas por los medios de comunicación, etc.

Para los alumnos que nos ocupan, alumnos terminando la educación secundaria, con muchas posibilidades por tanto de acceder a la formación universitaria, la situación ideal consiste en tener, por un lado, una información suficiente de la naturaleza de cada una de las profesiones, de la formación necesaria para acceder a ellas, de sus condicionamientos económicos y sociales, y de las condiciones del mercado de trabajo al que han de acceder en su momento, y, por otro, un esquema de valoración de

la información recibida, proveniente de sus experiencias académicas previas, del conocimiento de las propias inclinaciones y de las propias capacidades, etc.

Los alumnos toman decisiones de acuerdo a la información y los esquemas que tienen disponibles. Esas decisiones vienen condicionadas desde muy atrás en su vida escolar. Aquellos alumnos que hayan tenido una experiencia negativa con una materia tenderán a evitar ésta en el futuro. En este sentido, la experiencia escolar es muy determinante en sentido negativo.

La importancia de las experiencias escolares relativamente tempranas ha sido reconocida por Berryman (1983), quien ha desarrollado el término «Scientific/mathematic talent pool», que podríamos traducir como «la reserva de talentos científico-matemáticos». Esta reserva estaría formada por aquellos alumnos que en su día terminarán escogiendo una profesión ligada a la matemática, la ciencia o la técnica. Se trata de un concepto muy poderoso. En cada curso o nivel, debe existir una definición operativa distinta. Puede definirse desde el comienzo de la escolaridad. ¿Por qué se define esa reserva y no otra distinta? ¿Qué tienen de específico las matemáticas y las ciencias? Pues parece ser que las opciones profesionales se hacen por grandes áreas, y el mecanismo de exclusión es muy importante. Es decir, en muchas ocasiones los alumnos que están en últimos cursos de secundaria tienen más claras cuáles son las profesiones que rechazan que las que eligen. Y entre las profesiones más a menudo rechazadas están aquellas que implican conocimientos matemáticos o científicos.

Este concepto nos permite analizar la evolución de la «reserva», estudiando en cada etapa las características de sus componentes, cuántos y quiénes la abandonan o pasan a engrosarla, en qué momentos y por qué razones. Ese estudio permitiría formular las acciones a emprender para lograr que un mayor número de mujeres formasen parte de ese grupo. Según estableció Berryman, en Estados Unidos ese grupo ya había surgido para noveno grado (14 años aproximadamente), y ya estaba completo para el grado doce (17 años aproximadamente). Durante el grado trece todos los movimientos eran hacia afuera, no hacia dentro de la «reserva». Hilton y Lee (1988) señalan que prácticamente en todas las transiciones se producen pérdidas en este grupo, siendo las mayores las que se producen al final. Si ese mismo esquema se confirmase en España, implicaría la adopción de medidas muy distintas en los distintos niveles. Dado el esquema escolar español, el trabajo debiera comenzar antes de la primera transición importante, es decir, de la EGB al BUP. Porque naturalmente en el último curso de la educación secundaria esa reserva está ya prácticamente consolidada.

Maple y Stage (1991) han llevado a cabo un estudio longitudinal, utilizando el concepto de «reserva» que hemos mencionado. En un análisis causal de un modelo longitudinal estudiaron la influencia de algunas variables de «background» en la elección de carreras de matemáticas-ciencias para alumnos y alumnas de raza blanca y de raza negra. Las variables exógenas incluían la formación de los padres, locus de control, influencia de los padres en la toma de decisiones por parte de los estudiantes, la influencia de la escuela en la toma de decisiones de los alumnos, puntuaciones tipificadas de cinco tests de rendimiento, y actitudes hacia las matemáticas. Las variables endógenas eran un bloque de variables relativas a la formación recibida en secundaria, y como variable final la elección de carrera realizada por cada uno de los sujetos de análisis. Sin entrar a detallar los resultados obtenidos en esa investigación, resaltaremos lo más importante. En primer lugar, el campo de estudio señalado como elección preferida al comienzo de la educación secundaria tiene un gran poder de predicción de la elección realmente hecha años más tarde, al llegar a la universidad. Esto no hace más que resaltar la importancia de la intervención temprana si se desea incrementar la presencia de mujeres en el ámbito de las ciencias-técnicas.

Entre las mujeres blancas, el alto rendimiento en matemáticas-ciencias en la educación secundaria **NO** predice necesariamente la persistencia en lo que hemos llamado la «reserva de talentos científico-matemáticos». Eso quiere decir que hay otros factores que están influyendo en los intereses profesionales de las mujeres. Algunos de esos factores ya han sido mencionados, por ejemplo, los estereotipos

acerca de la masculinidad o femineidad de ciertas carreras, y la tendencia de todos los sujetos de adaptarse en mayor o menor grado a esos estereotipos.

También entre las mujeres negras la actitud hacia las matemáticas, al comienzo de la educación secundaria, era un buen predictor de la posterior elección de carrera de ciencias, de matemáticas o técnica.

Señalaremos para terminar este apartado, que las diferencias entre hombres y mujeres, una vez llegados al final de la educación secundaria, a la hora de elegir profesión, parecen mayores que las diferencias encontradas entre esos mismos individuos en función de la clase social. Ware y Lee (1988) por ejemplo encontraron que el estatus socioeconómico de la familia era un predictor positivo de la elección de carrera de ciencias sólo para los hombres y no para las mujeres. Esto parece indicar que el rechazo o al menos el desapego del ámbito de las matemáticas, la ciencia y la técnica está uniformemente distribuido entre las mujeres de todas las capas sociales.

FORMULACIÓN DEL PROBLEMA DE INVESTIGACIÓN

Vistos los antecedentes más importantes, el problema de investigación que nos ocupa, queda expresado en los siguientes términos: En 1987, en España, ¿Difieren los chicos y chicas, alumnos de último curso de secundaria en sus intereses profesionales? ¿Están modificadas las posibles diferencias inter-géneros por la clase social a la que pertenecen esos chicos y chicas?

Estas son las preguntas a las que en definitiva pretende aportar algún esbozo de respuesta esta investigación. Las hipótesis quedarían expresadas como:

H₁: El sexo determina diferencias en los intereses profesionales manifestados por los alumnos de último curso de educación secundaria.

H₂: La clase social determina diferencias en los intereses profesionales manifestados por los alumnos de últimos curso de educación secundaria.

H₃: La clase social modifica las diferencias en intereses profesionales entre chicos y chicas de último curso de educación secundaria.

VARIABLES INDEPENDIENTES

Las variables independientes son el sexo y la clase social. Si la variable sexo no presenta mayores complicaciones, la variable clase social es más compleja. Para empezar, ninguna de las variables incluidas en el banco de datos responde a esa etiqueta. Por otra parte, la determinación de la naturaleza de la variable «clase social» presenta complejidades de orden teórico. Sería preciso un estudio específico acerca de la naturaleza del «constructo clase social», sin olvidar la posibilidad de que esa discusión llevara a la conclusión de que las variables presentes en este banco de datos no miden o no reflejan adecuadamente lo que se entiende por dicha expresión. Por eso, en este estudio se parte de la posición más pragmática de considerar la expresión clase social como referida a unas características de los sujetos reflejadas en algunas de las variables recogidas. Con esto señalamos que no se pretende que la clase social se limita a aquello que reflejan estas variables.

De entre todas las variables se seleccionaron aquellas que mejor podían reflejar el estatus familiar. Por sus relaciones con aptitudes, actitudes y rendimientos, las variables que mejor reflejan el estatus socioeconómico o sociocultural familiar, son la profesión del padre (APROPSR), la estabilidad laboral del padre (ESTILAP), la situación laboral del padre (SILAPA), los estudios del padre (ESTPAD), los estudios de la madre (ESTMAD), y la situación laboral de la madre (SILAMA).

El hecho de que ninguna de estas variables represente perfectamente lo que entendemos como el estatus social de la familia nos llevó a planear algunos análisis que nos permitiesen crear una variable

que, de alguna forma, recogiese la información más importante contenida en el conjunto mencionado. Se trata de la variable CLUSTER.

VARIABLES DEPENDIENTES

Las variables dependientes son aquellas que reflejan los intereses profesionales de los sujetos analizados. Los intereses profesionales se midieron con el cuestionario CIPSA. Este cuestionario consta de 192 profesiones agrupadas en 12 grupos de dieciseis. Los nombres de las doce escalas o grupos, son: FÍSICO-QUÍMICA, DERECHO y LEGISLACIÓN, MEDICINA y SANIDAD, SERVICIOS, CIENCIAS HUMANAS, CIENCIAS BIOLÓGICAS, ACTIVIDADES LITERARIAS, PUBLICIDAD y COMUNICACIÓN, ARTES PLÁSTICAS y MÚSICA, ORGANIZACIÓN y MANDO, ENSEÑANZA, RELACIONES ECONÓMICAS Y EMPRESARIALES. El nombre de estas escalas no siempre refleja de forma precisa el contenido de las mismas².

Para cada profesión, los sujetos deben responder en tres variables, la valoración personal, con cuatro categorías (desconocida, indiferente, rechazada, elegida), y las valoraciones social y económica, con una escala de 1 a 4 cada una de ellas. Sólo la primera de dichas escalas es de interés en este estudio. Tal y como se presenta esta escala, no es fácil obtener una información global. Por ello se recodificaron las variables, de forma que a la indiferencia se le asignó un cero, a los rechazos un valor de -1, y a las elecciones de +1. El desconocimiento de una profesión se ha considerado como dato perdido. A continuación, para cada escala se obtuvo la suma de las puntuaciones asignadas a cada profesión, con lo que tenemos una escala posible desde -16 hasta +16. Los valores negativos significan rechazo, y los positivos aceptación.

Es preciso resaltar que lo que se mide no son las elecciones profesionales realmente realizadas por los sujetos, sino, como indica el nombre del cuestionario, sus intereses profesionales. Se trata de deseos o actitudes, y no siempre la elección profesional efectiva corresponderá con esas expresiones.

Por otra parte, es probable que muchos de los alumnos de último curso de secundaria tengan ya perfilada su elección profesional. Eso hace que elijan unas pocas profesiones y rechacen las demás. Por eso era lógico esperar valores mayoritariamente negativos en las escalas analizadas.

VARIABLES MODERADORAS

Además de las variables dependientes referidas a las elecciones profesionales, hay otro conjunto de variables endógenas que tienen un papel intermediario. Se trata de aquellas variables aptitudinales, actitudinales y de rendimiento escolar, a través de las cuales se ejerce el influjo de las variables independientes sobre las variables dependientes. El banco de datos incluye tres variables aptitudinales: Aptitud numérica, medida por el test DATNA, Razonamiento Verbal, medida por el test DATVR, y Razonamiento Abstracto, medida por el test DATAR. La batería DAT (Differential Aptitude Tests) fue construida inicialmente para medir varias aptitudes³ en alumnos comprendidos entre 8º y 12º grado de High School, equivalente a las edades de 14 a 18 años. Los tres tests que están incluidos en el banco de datos están ligados a funciones asociadas a la inteligencia general, y tiene por tanto sentido que

2 Se aconseja consultar el manual del cuestionario, en Fernández Seara, J. L. y Andrade, F. *Cuestionario de intereses profesionales* TEA Ediciones, Madrid 1983.

3 La batería completa consta de ocho tests, DAT-VR para Razonamiento Verbal, DAT-NA para Aptitud Numérica, DAT-AR para Razonamiento Abstracto, DAT-SR para Relaciones Espaciales, DAT-MR para Razonamiento Mecánico, DAT-CSA para Rapidez y precisión Perceptivas, DAT-LU-I para Uso del lenguaje-ortografía, y DAT-LU-II para Uso del lenguaje-Sintaxis.

hayan sido esos los tests de la batería elegidos para caracterizar las aptitudes de los alumnos de la muestra.

Las variables referidas a actitudes personales de los alumnos fueron operacionalizadas a través de cuatro escalas del cuestionario M.O.S.A.S. La escala L o Motivación de Logro consta de 20 preguntas del tipo verdadero o falso. Todas las preguntas sitúan al sujeto ante frases que se refieren a la disyuntiva de afrontar un riesgo para mejorar su situación o tratar de conservar lo que ya se posee. Cada enunciado tiene el valor de cero o de uno en función de si es congruente con el extremo positivo del continuo o no. La puntuación en la escala se obtiene sumando los puntos de cada uno de los 20 ítems. Las escalas E, Esperanza de éxito, P, Perseverancia, y A, autoestima, constan de 30, 30 y 20 ítems respectivamente, puntuándose de igual forma que la escala L.

DESCRIPCIÓN GENERAL DE LA MUESTRA

Los datos que aquí se analizan proceden de muestra para la evaluación de la reforma de las enseñanzas medias recogida por el CIDE.

La muestra está compuesta por 3.519 hombres y 3.746 mujeres.

Son alumnos de último curso de secundaria. Eso quiere decir que en todo el proceso educativo anterior se han ido produciendo deserciones que hacen que el grupo ahora analizado sea de alguna forma una élite entre los individuos de su misma edad. Se ha restringido el rango de muchas variables por el simple hecho de que estos alumnos son los que «han sobrevivido» en el sistema educativo. Esa es la razón por la que las relaciones entre muchas variables serán menos intensas por existir la mencionada reducción de rango. Por ese motivo es posible que la influencia de la clase social sobre los intereses profesionales sea más fuerte en la población general de lo que aquí aparece. Pero eso es algo que por definición no puede medirse, ya que aquellos sujetos que están fuera del sistema educativo a esta edad no pueden aspirar a muchas de las profesiones aquí descritas. Una solución consistiría en estudiar esas relaciones antes de comenzar la educación secundaria. Pero los intereses profesionales no están cristalizados a tan temprana edad. Como consecuencia de todo lo anterior, es preciso considerar con prudencia los resultados aquí presentados.

DESCRIPCIÓN DEL DISEÑO Y DE LAS HIPÓTESIS QUE SE CONTRASTAN

En nuestro caso tenemos un diseño univariado con dos variables independientes, (SEXO y «clase social» (CLUSTER) o SEXO y «estudios de la madre» (ESTMAD)) de efectos fijos, y las hipótesis que vamos a contrastar son las siguientes:

$$\begin{aligned} H_{01}: \mu_i &= \mu_r \\ H_{02}: \mu_j &= \mu_j \\ H_{03}: \mu_{ij} - \mu_{i'j} &= \mu_{ij'} - \mu_{i'j'} \end{aligned}$$

para todos los valores de i, i', j, j' siendo $i < i'$ y $j < j'$, y todas las $n(i,j) > 0$.

Las dos primeras hipótesis nulas se formulan para estudiar los efectos principales, mientras que la tercera es para estudiar la interacción. En otros diseños, el estudio de la interacción es bastante complejo, pero nosotros tenemos un diseño 5×2 o 7×2 , según la combinación de variables a que nos referimos. En ese caso, la interacción tiene una fácil interpretación. H_{03} propone que en todas las clases sociales, la diferencia en las medias de la variable dependiente entre los chicos y las chicas son constantes. Es decir, si no hay interacción, las diferencias entre chicos y chicas se mantienen constantes. Si hay interacción, al cambiar de clase social las diferencias entre chicos y chicas aumentan o

RESUMEN DE LOS RESULTADOS

	ESTMAD X SEXO			CLUSTER X SEXO		
	ESTMAD	SEXO	INTER.	CLUSTER	SEXO	INTER.
FISIQUI	6,41	208,70	1,64	7,42	403,80	0,09
	6,6813	1,6813	6,6813	4,6923	1,6923	4,6923
	0,0000	0,0000	0,1323	0,0000	0,0000	0,9871
DERLEGI	2,54	28,62	2,31	4,96	39,57	1,97
	6,68	1,6800	6,6800	4,6907	1,6907	4,6907
	0,0185	0,0000	0,0313	0,0005	0,0000	0,0959
SALUD	4,95	18,59	0,59	8,35	27,80	0,42
	6,6792	1,6792	6,6792	4,6902	1,6902	4,6902
	0,0000	0,0000	0,7386	0,0000	0,0000	0,7960
SERVICI	42,06	0,57	1,36	77,68	1,24	1,78
	6,6801	1,6801	6,6801	4,6910	1,6910	4,6910
	0,0000	0,4499	0,2253	0,0000	0,2648	0,1301
CIENHUMA	1,88	92,66	0,36	3,36	144,65	4,20
	6,6776	1,6776	6,6776	4,6883	1,6883	4,6883
	0,0807	0,0000	0,9047	0,0094	0,0000	0,0021
CIENBIO	1,78	5,88	1,26	2,22	4,94	1,32
	6,6754	1,6754	6,6754	4,6862	1,6862	4,6862
	0,0984	0,0154	0,2709	0,0642	0,0263	0,2602
LITERAR	0,37	10,13	1,01	0,40	12,72	0,10
	6,6627	1,6627	6,6627	4,6738	1,6738	4,6738
	0,8991	0,0015	0,4156	0,8059	0,0004	0,9816
PÚBLICO	3,40	7,44	1,85	3,65	3,20	2,13
	6,6607	1,6607	6,6607	4,6716	1,6716	4,6716
	0,0024	0,0064	0,0849	0,0057	0,0736	0,0743
PLASTMU	1,96	4,14	1,84	2,85	5,27	1,33
	6,6545	1,6545	6,6545	4,6653	1,6653	4,6653
	0,0672	0,0419	0,0870	0,0224	0,0217	0,2546
ORGMAND	4,05	2,89	0,51	3,57	2,94	0,67
	4,6616	1,6616	4,6616	6,6509	1,6509	6,6509
	0,0028	0,0890	0,7309	0,0016	0,0863	0,6715
ENSEÑAN	3,72	3,23	0,80	6,51	2,15	0,85
	6,6416	1,6416	6,06416	4,6519	1,6519	4,6519
	0,0011	0,0722	0,5664	0,0000	0,1425	0,4952
ECOEMPR	8,16	0,40	2,11	11,89	1,64	4,04
	6,6368	1,6368	6,6368	4,6472	1,6472	4,6472
	0,0000	0,5275	0,0489	0,0000	0,2002	0,002

F, grados de libertad, P(F)

disminuyen. Eso implica que si representamos gráficamente la relación entre clase social y variable dependiente por una parte para los chicos y por otra para las chicas, la no existencia de interacción nos daría dos curvas aproximadamente paralelas. Mientras que de existir interacción esas curvas llegarían a cruzarse en algún punto dado.

CONCLUSIONES

Relaciones entre la clase social y los intereses profesionales

A lo largo de los análisis que hemos realizado, obtenemos en una mayoría de los casos unas diferencias estadísticamente significativas en los intereses profesionales manifestados debidas a la variable clase social. Al estudiar el sentido o la tendencia de estas diferencias comprobamos que en un gran número de casos, esa relación es de tal forma que al aumentar el estatus social, se produce un aumento en el rechazo de las profesiones estudiadas. Y eso con independencia de si se trata de chicos o de chicas. Eso nos lleva a preguntarnos acerca de cómo responden los sujetos al cuestionario. Imaginemos dos alumnos que se encuentran en distintas etapas de la formación de sus intereses profesionales. En el primer caso, muchas de las profesiones no se conocen, o no se tiene una actitud definida hacia ellas. En el segundo caso la «vocación profesional» está claramente decidida o formada, y se identifica con un sólo elemento de una cierta escala. El primer sujeto tiene mayores probabilidades de rechazar menos profesiones que el segundo. Este último elegirá el elemento en cuestión, y con relación a los demás tendrá mayores probabilidades de rechazarlos, ya que no coinciden con su elección. **La hipótesis es por tanto que los individuos de las clases sociales más altas tienen unos intereses profesionales más cristalizados, llegando a ello por distintos mecanismos que ahora no interesa detallar.** Si esta hipótesis fuese válida, nos encontraríamos con que los individuos de clases sociales más altas darán mayor número de rechazos. Debemos pensar que elegir por exclusión también es elegir. Sería interesante poder probar esta hipótesis con un estudio específicamente diseñado al efecto. Una idea relacionada con esta hipótesis, hace referencia al número de opciones realizadas. Es decir, merecería la pena estudiar estas profesiones desde el punto de vista de su conocimiento o desconocimiento declarado por los sujetos, por una parte, y por otra, desde el punto de vista del número de opciones (tanto rechazo como elección) en oposición a la frecuencia de la declaración de indiferencia. Existe la posibilidad que alguno de estos otros enfoques nos llevasen a resultados relevantes.

Relación entre el sexo y los intereses profesionales

Seguramente la conclusión más clara a la que se puede llegar en este estudio hace referencia al influjo del sexo en los intereses profesionales. Hay una serie de escalas en las que hay una mayor valoración por parte de las chicas. Esas escalas son DERLEGIS, SALUD, CIENHUMA, LITERARI. Otra escala es con gran diferencia más preferida por los chicos que por las chicas. Se trata de FISIQUM. Otras dos tienen diferencias que están al borde de la significación estadística. En el caso de CIENBIOL, son los chicos los que aceptan ligeramente mejor que las chicas esta escala, siendo al contrario con la escala PLASTMUS. Por último hay cinco escalas, SERVICIO, PUBLICOM, ORGANMAND, ENSEANZ, y ECOEMPRES en las que no hay diferencias entre chicos y chicas. En estas cinco sí había efectos significativos debidos a la clase social. En dos escalas no hubo diferencias debidas a la clase social, CIENBIOL Y LITERARI. En otras dos, DERLEGIS y CIENHUMA, las diferencias están al borde de la significación estadística.

Con relación a la pregunta, ¿Qué influye más en los intereses profesionales, el sexo o la clase social?⁴. De forma general podemos decir que cuando existen los dos efectos siempre es mucho más acusado el de la variable SEXO que el de cualquiera de las otras dos variables estudiadas, «clase social» (CLUSTER) o «estudios de la madre» (ESTMAD). Cuando, como en las cinco escalas que hemos mencionado, no existen efectos debidos al sexo, los efectos debidos a la clase social puede decirse que responden a un esquema general ya mencionado en el apartado anterior. Las diferencias debidas al sexo son más específicas, y mucho más informativas. Y sobre todo porque tenemos alguna idea que nos permite dar un sentido a la aparición de esas diferencias, mientras que en el caso de la clase social hemos tenido que construir una hipótesis «ad hoc» para dar explicación a las diferencias encontradas.

De lo que no cabe la menor duda tampoco es que la clase social no establece ninguna modificación a las diferencias que encontramos entre los sexos cuando éstas existen. Es decir no existe, de forma general, interacción entre las dos variables. Este hecho debe ser interpretado a la luz de dos ideas. En primer lugar, no debemos olvidar que de alguna forma la muestra que hemos analizado es una muestra sesgada, ya que está compuesta por todos aquellos sujetos, de cualquier modalidad de educación secundaria, que han legado hasta el último curso de la misma. Seguramente ha habido ya una criba previa, por la que los sujetos que peor preparados estaban tuvieron que abandonar el sistema educativo en momentos anteriores. No son por tanto representativos de las clases baja, media-baja y media los sujetos asignados a dichos grupos. Puede que exista interacción entre las variables implicadas, pero tal vez es metodológicamente imposible descubrirla. La otra idea es que si esa criba no se modifica, es un dato cierto que con los sujetos «supervivientes» no hay interacción entre el sexo y la clase social, lo que quiere decir que si se pensase en cualquier campaña para tratar de aumentar el número de mujeres en las carreras y profesiones científico-técnicas, no sería preciso modificar la campaña en función de la clase social, ya que la distribución de las diferencias en intereses profesionales entre chicos y chicas es uniforme a lo largo de las distintas clases sociales.

Si hubiéramos de destacar algún dato por encima de los demás es el hecho de que la mayor diferencia entre chicas y chicos se produce justamente en la escala FISIQUM, mostrando las chicas un gran rechazo a las profesiones englobadas en ella. Esto, unido a las diferencias que se observaron en los análisis de las variables moderadoras hace que surjan muchos interrogantes en lo que se refiere a la explicación de los hechos, y que se sugieran posibles líneas de acción. Está claro que existe un alejamiento de las chicas con relación a las áreas técnico-científicas. Ese es un hecho que ya era conocido o intuitivo. Pero la evidencia empírica es aplastante. Hemos resaltado el hecho de la «universalidad» de esas diferencias con relación a las clases sociales, algo no tan conocido o no conocido de ningún modo. Pero queda todavía por delante la tarea de tratar de explicar detalladamente las causas de ese alejamiento, a través de otras investigaciones.

Por otra parte los análisis de covarianza que hemos realizado nos han demostrado que aunque alguna aptitud, como la medida por el test DATNA fundamentalmente, tienen alguna pequeña capacidad para explicar las distancias en intereses profesionales entre los dos sexos, esa capacidad de explicación era mínima, y nos obliga a concluir que ni las aptitudes, ni las actitudes son las variables que están provocando esa distinta actitud de los chicos y las chicas ante las profesiones, fundamentalmente ante las relacionadas con las ciencias, la técnica y las matemáticas. Lo que nos lleva a mantener que hay otras causas relacionadas con la asociación estereotípica de ciertas profesiones y actividades

4 Por supuesto se trata de una pregunta retórica. Por su naturaleza de atributos, de forma estricta ni una ni otra pueden 'influir' en los intereses profesionales. Con esta pregunta en realidad nos preguntamos acerca de qué información nos permite predecir mejor los intereses profesionales. En realidad la formación de los intereses profesionales está asociada a una serie de 'causas verdaderas' que se 'asocian' al sexo o a la clase social. Se trata de variables contextuales que aquí no se han recogido.

con cada uno de los sexos. No deja de ser significativo que incluso el instrumento empleado para medir los intereses profesionales podría achacársele fundamentalmente un cierto carácter sexista, por la forma en que se construyeron cada una de las escalas, y por el género asignado a ciertas profesiones.

CONSECUENCIAS DE ORDEN PRÁCTICO

No es el objeto de este estudio la propuesta de determinadas líneas de actuación administrativa o política. Pero es difícil resistir la tentación de resaltar algunas consecuencias de orden práctico que parecen derivarse de los resultados aquí obtenidos. A la sociedad le compete la tarea de tratar de reducir ese alejamiento para que la igualdad de oportunidades sea una realidad efectiva, y a los investigadores el tratar de explicar los hechos y si es posible determinar la mejor forma de evitar ese apartamiento de muchas alumnas de una de las áreas que están experimentando precisamente mayor desarrollo.

Sería preciso el desarrollo de una campaña institucional dirigida a las alumnas de secundaria o incluso a los últimos cursos de primaria, a sus padres y a sus profesores en la que se resaltase el hecho de que no existen profesiones o carreras a priori masculinas o femeninas. Romper esos prejuicios es la única forma de lograr equilibrar la presencia de los dos sexos en esas áreas. Debe conseguirse un cambio de actitud en las chicas hacia las matemáticas y las ciencias. Tal vez pudiese lograrse haciendo que conociesen más a fondo y sin prejuicios sexistas la verdadera naturaleza del trabajo científico. Todo esto supone campañas de carácter académico, programas de actuación que presenten las ciencias de forma atractiva, campañas de información acerca del papel de importantes mujeres en la ciencia (Ada Lovelace, Marie Curie, Barbara McClintock, etc.), campañas de información acerca de la utilidad de la ciencia en el desarrollo social, especialmente dirigido a las mujeres etc.

PROPUESTA DE POSTERIORES ESTUDIOS

Para terminar este trabajo vamos a formular una serie de preguntas para las que la respuesta supondría llevar a cabo posteriores investigaciones. En primer lugar, ¿Responden chicos y chicas de forma distinta al cuestionario? ¿Tienen los chicos una actitud más cristalizada respecto a las profesiones que las chicas? ¿Existen diferencias en la forma de responder al cuestionario entre las distintas clases sociales? ¿Hay diferencias en el nivel de conocimiento de las distintas profesiones entre chicos y chicas? ¿Existen esas diferencias entre las clases sociales? En otro orden de cosas, ¿Puede establecerse un modelo causal que explique algo mejor la cadena de influencia o de concomitancia entre sexo y clase social respecto a los intereses profesionales? ¿Existe sesgo en los ítems de los instrumentos contruidos para medir aptitudes y rendimientos, con respecto a las chicas?

Apartándonos de los datos utilizados hasta el momento, ¿Es el «autoconcepto» una variable importante a la hora de explicar las diferencias en intereses profesionales entre chicos y chicas? ¿Es modificable de algún modo el «autoconcepto» a través de intervención externa? Es obvio que la respuesta a estas últimas cuestiones sería importantísima si se tiene en perspectiva una futura intervención institucional para reducir las diferencias mencionadas.

REFERENCIAS BIBLIOGRÁFICAS

- ÁLVARO PAGE, Mariano y otros (1988): *Estudio de evaluación externa de las enseñanzas medias (I)*. C.I.D.E., Madrid.
- BERRYMAN, S. E. (1983): *Who will do science? Minority and female attainment of science and mathematics degress: Trends and causes*. The Rockefeller Foundation, New York.

- HILTON, T. L. y LEE, V. E. (1988): *Student interes and persistence in science: Changes in the educational Pipeline in the last decade*. Journal of Higher Education, 59, pp. 510-526.
- INSTITUTO DE LA MUJER (1987): *Situación social de la mujer en España*.
- MAPLE, S. A. y STAGE, F. K. (1991): *Influences on the choice of Math/Science major by gender and ethnicity*. American Educational Research Journal, vol. 28, nº. 1, pp. 37-60.
- WARE, N. C. y LEE, V. C. (1988): *Sex differences in choice of college science majors*. American Educational Research Journal, 25, pp. 593-614.

APROXIMACIÓN PSICOFISIOLÓGICA AL ESTUDIO DEL PROCESO ATENCIONAL EN NIÑOS CON SÍNDROME DE DOWN

Francisco A. García Sánchez

Universidad de Murcia

José M. Martínez Selva

Facultad de Psicología. Universidad de Murcia.

INTRODUCCIÓN

En los últimos años, asociado al desarrollo de nuevos paradigmas experimentales, ha crecido el interés por estudiar los procesos perceptuales, atencionales y de memoria en niños con Síndrome de Down (SD) (Wagner, Ganiban y Cicchetti, 1990). Un ejemplo de ello es el presente trabajo, en el que estudiamos los procesos atencionales básicos en el niño con SD valiéndonos de una técnica y de una metodología psicofisiológica que nos proporciona una medida objetiva de los mismos. En concreto hemos estudiado la respuesta de orientación (RO), su habituación y su reaparición ante el cambio estimular en una muestra de niños con SD y en un grupo de control emparejado en sexo y edad cronológica.

La RO se considera una forma de atención fásica o pasajera que constituye la base de otras formas más complejas de atención. Se trata de la primera reacción del organismo ante cualquier estímulo nuevo o significativo y se relaciona con la formación de modelos neuronales, o huellas mnésicas, de esos estímulos. Los diferentes componentes del RO y sus parámetros nos informan acerca de un tipo de proceso atencional que tiene importantes implicaciones para el aprendizaje y el desarrollo cognitivo. El desencadenamiento de la RO por los estímulos indica el nivel de conciencia del sujeto ante los cambios ambientales (Sokolov, 1963). Este nivel de conciencia queda reflejado en los parámetros psicofisiológicos vegetativos de la RO, tales como la magnitud de la respuesta de conductancia de la piel, especialmente de la RO inicial, el descenso en el volumen de pulso periférico (VPP) y la desaceleración cardíaca (Barry, 1982, 1990). Por su parte, la habituación de la RO indica la velocidad de formación, en la corteza cerebral, de esas huellas mnésicas, o «modelos neuronales del estímulo» en la terminología de Sokolov, y la capacidad del sujeto para ignorar información irrelevante (Sokolov, 1963).

Los trabajos que han empleado medidas electroencefalográficas de la atención y de la RO han encontrado potenciales evocados de mayor amplitud ante la estimulación repetida en niños con SD que en otros grupos de niños con o sin retraso mental. Este resultado suele interpretarse como la consecuencia de un déficit en procesos inhibitorios y, más concretamente, como una dificultad del sistema nervioso central para habituar sus respuestas bioeléctricas ante la estimulación. Normalmente, los

procesos neuronales inhibitorios suprimen selectiva y específicamente las respuestas a estímulos irrelevantes o predecibles. La responsividad neuronal de los sujetos con SD podría verse incrementada por una ausencia de procesos inhibitorios, por una limitada selección de los estímulos y por una limitada especificidad de sus respuestas a los cambios ambientales (Courchesne, 1988; Galbraith, 1986). Cualquiera de estas alteraciones dañaría sus procesos atencionales y contribuiría a una pobre realización en un amplio abanico de actividades mentales.

Por el contrario, la literatura acerca de la habituación de la RO medida a través de índices vegetativos es bastante escasa para sujetos con SD y los resultados obtenidos no son concluyentes. La mayoría de las veces los niños con SD se incluyen en grupos de niños con retraso mental de etiología diversa, lo que dificulta la comparación de los trabajos. En general, cuando se observan diferencias, entre los grupos de retraso mental y de control, parecen depender de la fuerza de la RO inicial, con la habituación más lenta asociada a las respuestas mayores (Clausen y Sersen, 1983). Algunos trabajos han encontrado una menor responsividad en los sujetos con SD, lo que se interpreta como una hiporresponsividad de su actividad autonómica (Clausen, Lidsky y Sersen, 1976; Wallace y Fehr, 1970). Esta menor responsividad, acompañada por RO de baja amplitud a los estímulos ambientales, podría indicar una falta de conciencia en muchas situaciones diferentes y un importante déficit en la formación de huellas mnésicas. La ausencia de un mecanismo atencional básico interferiría en los estados iniciales del proceso de memoria.

Nuestro propósito fue estudiar la actividad electrodérmica (AED) y cardiovascular en niños con SD ante estímulos de moderada intensidad (80dB), con un paradigma simple de habituación y cambio estimular. Esperábamos encontrar menor responsividad vegetativa y, consecuentemente, una tasa de habituación más rápida, en los niños con SD que en el grupo de control.

MÉTODO

Sujetos

De un total de 42 niños, se seleccionaron 19 niños con SD (8 niños y 11 niñas), de edades comprendidas entre 7 y 13 años (media= 9.8, $dt= 1.11$), y otros 19 niños como grupo de control (hermanos de los niños con SD y emparejados uno a uno en sexo y edad cronológica). Todos los niños con SD recibían asistencia en la Asociación para el Tratamiento Precoz del Niño con Síndrome de Down (ASSIDO, Murcia). Su CI medio fue de 61.21 ($dt= 10.38$, rango= 39-82).

Instrumentos y procedimiento

El registro de la AED fue bipolar por medio de un amplificador de conductancia de la piel (LA76441) y unos electrodos de Ag/AgCl y 7mm de diámetro (LA76616) colocados sobre las eminencias tenar e hipotenar de la mano izquierda. Como medio de contacto se empleó un biogel 0.05 mol NaCl. El VPP y la tasa cardíaca se registraron a través de un transductor piezoeléctrico (LA76605) colocado sobre la falange distal del dedo pulgar de la mano izquierda, cuya señal fue amplificada por un monitor de pulso sanguíneo (LA76405).

Después de un periodo de línea base de 5 minutos, se presentaron 11 estímulos auditivos: 10 tonos de 2 seg.-80 dB-1000 Hz seguidos de un estímulo final de 2 seg.-80 dB-500 Hz (cambio estimular). El intervalo entre estímulos varió aleatoriamente entre 20 y 45 segundos (media= 30 seg.). En algunos niños con SD este intervalo se prolongó debido a la aparición de movimientos corporales o artefactos en sus registros fisiológicos que obligó a demorar su presentación unos segundos (< 10 seg.). Todo el procedimiento experimental duraba aproximadamente 15 minutos.

Identificación y cuantificación de respuestas

Dentro de la actividad electrodérmica analizamos:

— Magnitud de las respuestas específicas de conductancia de la piel ante los estímulos (primera respuesta que aparecía 1-5 segundos después de la presentación de cada estímulo).

— Habitación medida por el número de respuestas específicas antes de alcanzar el criterio de habitación, establecido en dos ensayos consecutivos sin respuesta.

— Frecuencia de respuestas, definida como el número de respuestas inespecíficas de conductancia de la piel por minuto. Medida durante los 5 minutos de línea base y durante la serie estimular.

— Niveles de conductancia de la piel, promediados para la línea base (se medían al principio y al final de la misma) y la serie estimular (se medían después del tercer y séptimo estímulo).

Dentro de la actividad cardíaca analizamos:

— Porcentaje de cambio del VPP, calculado entre la amplitud promedio (en milímetros) del VPP en los 4 segundos previos a la presentación de cada estímulo y la amplitud promedio del VPP en los 8 segundos posteriores al mismo.

— Diferencia en frecuencia cardíaca (FC) entre la FC media en los 4 segundos previos a la presentación de cada estímulo y la FC media en los 8 segundos posteriores al mismo.

RESULTADOS Y CONCLUSIONES

Los resultados obtenidos en los análisis de varianza uni— y multivariados aplicados a las distintas variables dependientes indican una menor responsividad electrodérmica en los niños con SD ante todos los estímulos ($F_{(1,31)} = 8.43, p=.007$), incluyendo la respuesta al cambio estimular ($F_{(1,31)} = 5.98, p=.02$), lo que apoya resultados previos de Clausen y cols. (1976), Fenz y McCabe (1971) y Wallace y Fehr (1970). Las variables cardiovasculares no mostraron diferencias entre los dos grupos, lo cual está parcialmente de acuerdo con la literatura existente ya que, si bien se han encontrado mayores desaceleraciones en niños con retraso mental no se han encontrado diferencias entre los grupos experimentales en VPP (Clausen y Sersen, 1983). Nuestros resultados apoyan dos de las conclusiones de la revisión de Clausen y Sersen: no se verifica un patrón organizado de respuestas autonómicas en sujetos con retraso mental y la respuesta electrodérmica aparece como la medida más sensible a las diferencias entre los grupos.

Dentro del marco de la teoría de la RO, nuestros resultados indican la existencia de importantes problemas en el desencadenamiento y habitación de la RO en los niños con SD:

1) Sus ROs son más pequeñas desde el primer estímulo. Esto sugiere una falta de los recursos atencionales iniciales necesarios para procesar el estímulo nuevo y para construir su «modelo neuronal» o huella mnésica.

2) Se produce un descenso brusco en la amplitud de la RO desde el segundo ensayo estimular en adelante (ver Figura 1). La habitación es una forma de aprendizaje que se manifiesta como un descenso progresivo en la amplitud de la respuesta ante la estimulación repetida. Las diferencias en la curva de habitación entre los dos grupos de niños estudiados son significativas cuando consideramos separadamente la RO inicial y el resto de respuestas a los estímulos. Como se observa en la Figura 1, los niños con SD prácticamente dejan de dar ROs a partir del segundo estímulo. Por su parte, los niños del grupo de control muestran un descenso progresivo en sus ROs, indicando el desarrollo de un proceso central de habitación ante un estímulo no-significativo de moderada intensidad. El hecho de que las ROs en los niños con SD no muestren el usual descenso pausado de la curva de habitación puede ser interpretado como un fallo en el formación de las huellas mnésicas asociadas al proceso de habitación, que impedirá futuros análisis de la relevancia de tales estímulos.

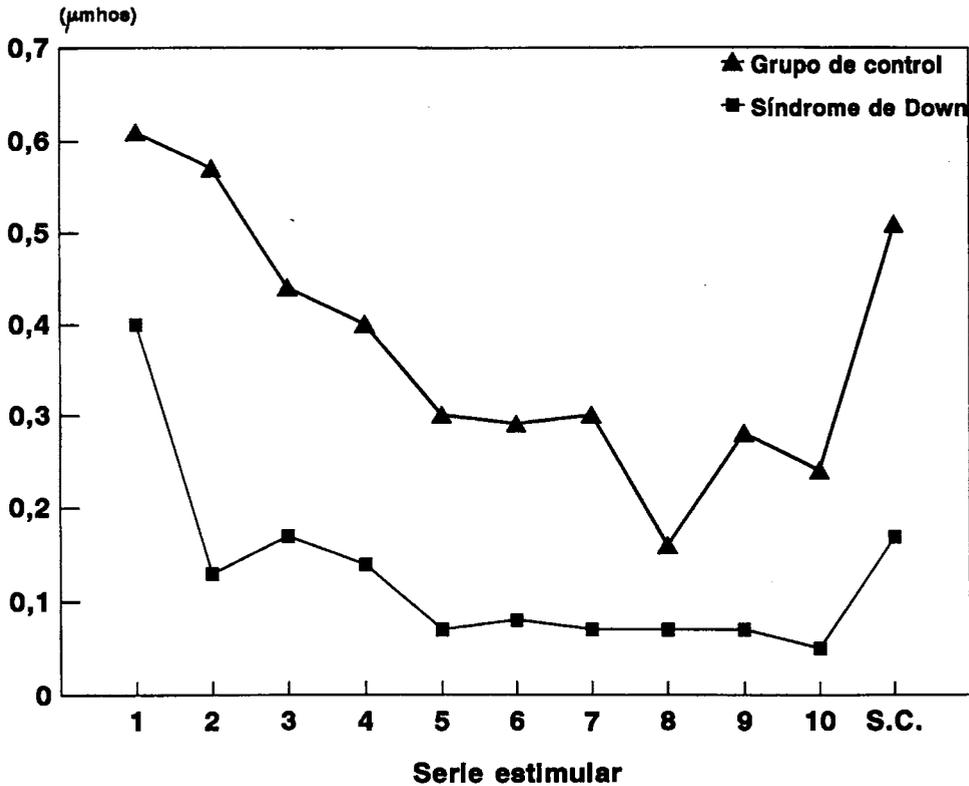


Figura 1
 Magnitud de las respuestas de conductancia de la piel a cada estímulo
 (S.C.= cambio estimular).

3) Posiblemente como una consecuencia de esa alteración en los procesos de habituación, la respuesta al cambio estimular fue menor en los niños con SD. Ya que la huella mnésica o «modelo neuronal» del estímulo no fue adecuadamente formada, el cambio estimular no fue detectado como tal, mientras que los niños del grupo de control si lo detectaron.

Los resultados de la RO electrodérmica en niños con SD describen una conducta atencional diferente a la que se infiere de los estudios con potenciales evocados. Las mayores amplitudes mostradas por estos niños en sus potenciales evocados y la no-habituación de sus respuestas, han sido interpretadas como un déficit en sus procesos inhibitorios (Courschene, 1988; Galbraith, 1986). Nuestros resultados indican un sistema nervioso hiporresponsivo. Sin embargo, estas diferentes aproximaciones pueden ser compatibles. La ausencia de selectividad ante los estímulos, que perjudica la adecuada discriminación entre ellos, es el déficit atencional encontrado en los estudios de la respuesta fisiología atencional en niños con SD. Si esos déficits atencionales son debidos a una falta de inhibición o a respuestas atencionales más pequeñas desde la primera presentación del estímulo es una cuestión a explorar en futuras investigaciones.

BIBLIOGRAFÍA

- BARRY, R. J. (1982): Novelty and significance effects in the fractionation of phasic OR measures: A synthesis with traditional OR theory. *Psychophysiology*, 19, 28-35.
- BARRY, R. J. (1990): The orienting response: Stimulus factors and response measures. *The Pavlovian Journal of Biological Sciences*, 25, 93-103.
- CLAUSEN, J. y SERSEN, E. A. (1983): The orienting response and intellectual retardation. En D. Siddle (Ed.): *Orienting and habituation: Perspectives in human research*, (pp. 505-522). Chichester: Wiley.
- CLAUSEN, J., LIDSKY, A., and SERSEN, E. A. (1976): Measurement of autonomic functions in mental deficiency. En R. Karrer (Ed.): *Developmental psychophysiology of mental retardation*, (pp. 39-91). Springfield, Ill: Charles C. Thomas.
- COURCHESNE, E. (1988): Physioanatomical considerations in Down syndrome. En L. Nadel (Ed.): *The psychobiology of Down syndrome* (pp. 291-313). Cambridge, MA: MIT Press.
- FENZ, W. D. and McCABE, M. W. (1971): Habituation of the GSR to tones in retarded children and nonretarded subjects. *American Journal of Mental Deficiency*, 75, 470-473.
- GALBRAITH, G. C. (1986): Unique EEG and evoked response patterns in Down syndrome individuals. En C. J. Epstein (Ed.): *The neurobiology of Down syndrome*, (pp. 109-119). Nueva York: Raven Press.
- SOKOLOV, E. N. (1963): *Perception and the conditioned reflex*. Londres: Pergamon Press.
- WAGNER, S.; GANIBAN, J. M. y CICCHETTI, D. (1990): Attention, memory, and perception in infants with Down syndrome: A review and commentary. En D. CICCHETTI and M. BEEGLY (Eds.): *Children with Down syndrome. A developmental perspective* (pp. 147-179). Nueva York: Cambridge University Press.
- WALLACE, R. M. and FEHR, F. S. (1970): Heart rate, skin resistance, and reaction time of mongoloid and normal children under baseline and distraction conditions. *Psychophysiology*, 6, 722-731.

COMUNICACIONES A LA PONENCIA II

CRITERIOS Y PROCEDIMIENTOS PARA LA SELECCIÓN DE INDICADORES DE RENDIMIENTO EN EVALUACIÓN INSTITUCIONAL UNIVERSITARIA

J. M. Osoro Sierra y L. Salvador Blanco
I.C.E. - Universidad de Cantabria

De forma tradicional las autoridades educativas e incluso los propios centros realizan considerables esfuerzos para la obtención de datos sobre distintos aspectos del funcionamiento de las instituciones todo ello con distintos propósitos en función de las distintas audiencias implicadas. No podemos considerar —ni pretendemos—, que todos los datos deban ser presentados en forma de indicadores de rendimiento. Ahora bien, siempre que se opte por esta vía, será necesario el establecimiento de algunos criterios de selección y dar prioridad a aquellos datos que permitan describir y caracterizar al sistema (total o parcial) de cara a la mejora de la práctica educativa.

La definición de indicador que adoptemos será la que determine los criterios de selección. No existe un acuerdo generalizado sobre la definición de indicador (distintas audiencias tienen distintos planteamientos), ni sobre un modelo que pueda servir para encuadrar el proceso educativo, sin embargo, existe el acuerdo unánime de que los indicadores necesitan ser manejables, precisos y útiles (Smith: 1988, p. 488). Conforme a estos aspectos podemos establecer cuáles son los criterios en la selección de indicadores (Wyatt, T. et al.: 1989, p. 65):

1. **Importancia/uso:** valor de la información para el desarrollo político, el interés de la audiencia y la rendición de cuentas.
2. **Calidad técnica:** basada en la validez del contexto y en la validez y fiabilidad de la información.
3. **Fiabilidad:** en función de la aceptación de una definición común, del cuidado en la recogida de datos, del costo de la recogida de datos, del análisis e informe y de la simplicidad de la información.

Debemos de tener en cuenta que la selección de criterios se desarrolla en un marco subjetivo. La elección está influenciada por una serie de valores y juicios cualitativos que requieren un alto consenso entre los implicados. En otras palabras **el diseño de indicadores debe servir para los objetivos establecidos de antemano y no para otros usos que no fueron previamente acordados**. Por tanto, son aspectos esenciales en la selección y construcción de indicadores el nivel de consenso logrado entre los implicados y la determinación del objetivo que se persigue con su utilización.

Desde un punto de vista general, McVicar, M. (1990) señala que la revisión y medición de la calidad universitaria es ejercitada por un grupo muy amplio de individuos, grupos y organizaciones, incluyendo estudiantes, personal, empleados, cuerpos profesionales, examinadores externos, inspecto-

res gubernamentales y otros. De todos ellos unos ejercen una influencia directa, por ejemplo, en la acreditación personal; otros tienen una influencia indirecta como, por ejemplo, el mercado. Sólo algunos son parte del proceso formal para asegurar la calidad.

A través de todo lo expuesto anteriormente, podemos señalar las notas y requisitos fundamentales del concepto de indicador:

1. Se plantea como una unidad de información, necesaria para ver el proceso de desarrollo (en el ámbito que sea) de la institución universitaria.
2. La información obtenida puede y debe ser tanto de índole cuantitativa como cualitativa.
3. Deben de estar ligados a ciertos criterios de calidad.
4. Es necesario en todo proceso evaluativo realizado a través de indicadores:
 - a) Identificar el tipo de indicador: input/proceso/producto.
 - b) Definir las funciones del indicador: cantidad/efectividad o calidad/eficacia.
 - c) Definir el centro hacia donde se dirige el indicador: personal/estudiantes/recursos...
5. Debe de haber un nivel de acuerdo previo entre los implicados.
6. La utilización de indicadores de rendimiento puede hacerse a través de evaluaciones externas, autoevaluaciones y sistemas de expertos (peer review).

Una vez que hemos definido los requisitos previos que deben de cumplir los indicadores, señalaremos los criterios de selección de los mismos. Cave, et al. (1988) señalan siete criterios técnicos que deben reunir todos los indicadores:

1. *Tipo de Indicador*: Entrada-Proceso-Resultado.
2. *Relevancia*: Está relacionado con los objetivos y misión de la institución o departamento.
3. *Claridad*: Está definido claramente, no es ambiguo.
4. *No Manipulable*.
5. *Costo de recogida de los datos y facilidad para comparar los datos*.

CRITERIOS TÉCNICOS

APLICACIÓN A LOS INDICADORES DE RENDIMIENTO

El indicador de rendimiento debe ser:

- * Relevante para los objetivos de investigación de la institución o departamento.
- * Definido claramente, comprensible para una amplia audiencia y abierto sólo a una interpretación.
- * Dificultad o imposibilidad de que sea manipulado
- * Capaz de ser relacionado con el nivel de análisis requerido.
- * Aceptado por los clientes que van a ser evaluados.
- * Capaz de ser clasificado en medidas de entradas-resultados-productividad o calidad.
- * Basado en datos fáciles de recoger en términos de: tiempo, costo y experiencia.
- * Verificable.

El indicador de rendimiento es:

- * Relevante o no.
- * Concreto u ambiguo.
- * Manipulable o no.
- * Capaz de relacionarse con el nivel del departamento o no.
- * Clasificado en una escala de 6 puntos (Muy de acuerdo-Nada)
- * Clasificado en una escala de 6 puntos. (Muy de acuerdo-Nada).
- * Aceptable en términos de tiempo, costo o experiencia en la recogida de datos.
- * Verificable o no.

6. *Nivel de situación*: Departamento, Disciplina, Profesor.

7. *Relación con otros indicadores*.

A estos siete criterios técnicos se les pueden añadir otros criterios, como son:

Coherencia: Equilibrio entre indicadores de docencia, investigación y servicios.

Confiabilidad: Válido y fiable en otras instituciones o departamentos.

Durabilidad: Válido y fiable a través del tiempo.

Tognolini, J. (1991, p. 162) nos presenta una serie de requisitos técnicos que deberán cumplir los indicadores de rendimiento y una propuesta de aplicación práctica de los mismos:

Desde otro punto de vista, Dochy, et al. (1990), a partir de su trabajo con indicadores, realizan una propuesta de construcción de los mismos basándose en tres aspectos:

1. Las funciones de la Universidad: Docencia-Investigación-Servicios.
2. El tipo de indicador: Entrada-Proceso-Resultado
3. La operativización de los indicadores en variables.
4. El tipo de medida: Cuantitativa-Cualitativa.

Nuestra metodología o propuesta de trabajo para la selección de indicadores se basada en la consideración y distinción entre los siguientes conceptos:

a. **OBJETIVOS**: Determinación del aspecto concreto que es motivo de evaluación, considerado por los implicados como un objetivo de calidad (infraestructuras, programas,...)

b. **DIMENSIONES**: Definición de los distintos elementos que componen el objetivo.

c. **INDICADORES**: Definición operativa de las dimensiones señaladas.

d. **VARIABLES**: Traslado de los indicadores a un determinado número de características específicas. Es el último nivel de concreción y operativización de los indicadores escogidos.

Teniendo en cuenta los criterios planteados en las páginas anteriores, pensamos que los indicadores de calidad pueden ser un instrumento válido para la evaluación de las instituciones universitarias. Planteando el concepto de indicador como información parcial o total del sistema, estos estarán encaminados a proporcionar datos sobre:

* Información sobre las entradas: lo que la institución ha hecho para transformar los recursos financieros en programas concretos de docencia, investigación y gestión.

* Información sobre la dedicación: describe cómo los diversos factores de entrada se han dedicado al logro de los objetivos.

* Información sobre el proceso: determina el modelo de producción escogido para alcanzar los resultados.

* Información sobre la utilización: cómo se usan los factores de entrada en el proceso.

* Información sobre el funcionamiento: informa sobre el proceso pero definiendo las características más importantes en función de su influencia en los resultados.

* Información sobre la eficiencia: valora la relación entre los recursos puestos a disposición de la institución y los resultados obtenidos.

REFERENCIAS BIBLIOGRÁFICAS

CAVE, et al. (1988): *The Use of Performance Indicators in Higher Education: A Critical Analysis of Developing Practice*. London: Jessica-Kingsley.

DOCHY, F., SEGERS, M., WIJNEN, W. (Eds.) (1990): *Management Information and Performance Indicators in Higher Education: an International Issue*. Assen/Maastricht: Van Gorcum.

GOEDEGEEBUURE, L., MAASSEN, P., WESTERHEIJDEN, D. (Eds.) (1990): *Peer Review and Performance Indicators*. Utrecht: Lemma.

- KELLS, H. R. (1990): *The Development of Performance Indicators for Higher Education. A Compendium for Eleven Countries*. París: OECD.
- McVICAR, M. (1990): Quality Assessment in British Higher Education. En GOEDEGEBUURE, et al: *Peer Review and Performance Indicators*. Utrech: Lemma.
- SMITH, M. S. (1988): Educational Indicators. *Phi Delta Kappan*, marzo.
- TOGNOLINI, J. (1991): Performance Indicators in Higher Education: Some Practical Problems and Issues. En HEWTON, J.: *Performance Indicators in Education: What Can They Tell Us?* Papers From the Third National Conference. Camberra: Australian Conference of Directors-General of Education, 159-171.
- WYATT, T., RUBY, A., NORTON, S., DAVIES B., SHRUBB, S. (1989): *Reporting on Educational Progress: Performance Indicators in Education. A report to the Conference of Directors-General of Education*. Sydney.

UN ESTUDIO SOBRE INDICADORES ESTRUCTURALES Y CALIDAD PERCIBIDA EN CURSOS DE POSTGRADO

A. Pérez Carbonell, J. M. Jornet, J. González Such y M. R. Ferrández
Universitat de València¹

1. INTRODUCCIÓN

La Universitat de València ha puesto en marcha diversos estudios encaminados a diseñar un Sistema de Evaluación de Cursos de Postgrado. Dentro de esta línea de investigación, se prevé la extracción de información que pudiera coadyuvar a la mejora de estos Cursos, de forma que la Evaluación no tenga exclusivamente la finalidad de rendición de cuentas prevista en la Normativa de la misma Universitat, sino que ésta se articule en un conjunto organizacional orientado a la mejora de las características estructurales y la funcionalidad de estos Cursos (Pérez Carbonell et al., 1993).

2. OBJETIVOS

En esta línea, se han considerado, en esta primera aproximación diversos tipos de indicadores estructurales, descriptivos de la forma en que se organiza cada Curso. Sí bien alguno de ellos ha sido considerado en otros contextos —análisis institucionales, evaluación de programas de larga duración en Educación Superior, etc.— y con otros objetivos —asignación de recursos, costes, etc.— (Goedegebuure et al., 1990; Johnes y Taylor, 1990; Dochy et al., 1990; Cave et al., 1991; Mora, 1991) han tenido menor repercusión en valoraciones de programas de corta duración. No obstante, su consideración en referencia a la calidad percibida por los «clientes» directos de este tipo de Cursos puede ofrecer información que oriente líneas de recomendación para su organización. Así, el Objetivo Principal de este estudio es analizar la dependencia de indicadores estructurales y la calidad percibida del Curso. Objetivo implicado será establecer si existen relaciones entre tales indicadores y la percepción de diversos aspectos globalizadores de la calidad y en consecuencia acercarnos al valor diferencial de este tipo de indicadores.

1 Avd. Blasco Ibáñez, 21, 46010-VALENCIA. Teléfono (96) 3864420/30.

2 Para extraer una variable síntesis de Logro de Objetivos se ha tomado la media de las respuestas a los ítems de cada Curso, teniendo en cuenta que todos se basaban en la misma Escala de puntuación (1 a 5) y estaban formulados con la misma dirección de respuesta.

3. METODOLOGÍA

3.1. Muestra

De los cuarenta Cursos de Postgrado que se ofertaban desde la Universitat de València durante el curso académico 91/92 se hizo una distribución en cinco áreas y se eligieron aleatoriamente 16 Cursos. El tanto por cien de encuestas recogidas en cada una de las áreas respecto al total de encuestas recogidas (376) fueron: en el área de Experimental un 14.10% , en el área de Humanidades un 13.56% , en el área de Psicopedagogía un 7.71% , en el área de la Salud un 34.84% y en el área de Sociales un 29.79%.

3.2. Variables

Los indicadores estructurales que se han tenido en cuenta son los siguientes:

Indicadores absolutos: a) Número de Horas Total (Número de horas totales de cada programa), b) Número Total de Materias (Número de materias que oferta cada programa), c) Número Total de Profesores (Número de profesores con docencia en el programa), d) Número Total de Alumnos (Número de alumnos matriculados en cada programa), y e) Coste Total (Fijado en el presupuesto que se recoge en el proyecto de cada programa).

Indicadores de Razón: a) Razón del Número de Horas por el Número de Materias (Promedio de horas destinadas a cada materia), b) Razón del Número de Horas por el Número de Profesores (Promedio de horas impartidas por cada profesor), c) Razón del Número de Materias por el Número de Profesores (Promedio de materias impartidas por profesor), d) Coste por Alumno (Promedio del coste del Curso por estudiantes), y e) Coste por Hora (Promedio del coste de la hora de cada programa).

La calidad percibida por los estudiantes se ha medido a partir de un Cuestionario global en el que se recaba información acerca de los siguientes aspectos —ver Tabla 1—: Aspectos Generales de Organización, cuestiones referidas al Profesorado que ha impartido los Cursos, logro de Objetivos del Curso y Utilidad y especificidad de Contenidos. El Cuestionario consta de un total de 21 ítems comunes y, 8 ítems relativos al logro de objetivos² que fueron especificados en relación con el Coordinador y profesorado de cada Curso.

3.3. Análisis

En este trabajo se presentan los resultados correspondientes a un primer acercamiento al problema. De esta manera, se informa del estudio, realizado sobre el total de la muestra, relativo a:

a) Análisis de Dimensiones Canónicas³ cuyo objeto ha sido sintetizar el esquema de relaciones entre Indicadores y Componentes Evaluativos. Así, se han realizado dos estudios canónicos: 1. Indicadores Absolutos/Componentes Evaluativos, y 2. Indicadores de Razón/Componentes Evaluativos.

b) Análisis Univariado de Diferencias significativas en cada uno de los ítems del Cuestionario de Evaluación entre los grupos formados a partir de la Mediana⁴ de cada indicador, mediante la prueba U de Mann-Whitney.

3 Los análisis se han realizado con el programa 6M de la versión Pc-90 de BMDP.

4 Los análisis se han realizado mediante el programa NPAR de la versión 5.0 del SPSS.

TABLA 1
SÍNTESIS DE ÍTEMS DEL CUESTIONARIO DE ESTUDIANTES REFERENTES A ASPECTOS DE ORGANIZACIÓN(AO) Y DESARROLLO DEL CURSO (DC)

CÓDIGO	PREGUNTA
AO 1	La información facilitada acerca de la organización del Curso/Máster (fechas, horarios, lugares de realización, inscripción, etc.) ha sido adecuada.
AO 2	En general, la organización del Curso/Master (fechas de realización, horarios elegidos, etc.) ha sido adecuada.
AO 3	La infraestructura de aulas disponible ha sido adecuada.
AO 4	Los materiales disponibles dentro de las aulas han sido adecuados.
AO 5	Se ha cumplido el calendario y horarios de sesiones previstos.
AO 6	Se han realizado todos los cursos previstos.
AO 7	El tiempo destinado en general a cada curso es: Suficiente, Excesivo o Escaso.
AO 8	Creo que podrían mejorarse algunos aspectos de la organización, indica cuáles.
DC 1	El profesorado ha planteado los objetivos del curso con claridad.
DC 2	Se ha podido apreciar una buena coordinación entre los profesores que han impartido el curso.
DC 3	La actuación de los profesores que lo han impartido ha sido adecuada.
DC 4	En general, los profesores no pertenecientes a la Universitat de València han realizado una aportación de interés para este curso.
DC 5	En el caso en que estés en desacuerdo con la actuación de algún profesor en concreto, por favor, indícanos el código del mismo, el de la materia que imparte y, finalmente, señala brevemente los motivos.
DC 6	Los materiales de trabajo (documentación, bibliografía, etc.) que se ofrecen durante el curso son adecuados.
DC 7	Los contenidos presentados durante el Curso son interesantes.
DC 8	Los contenidos presentados a lo largo del Curso son realmente útiles para mi ejercicio profesional.
DC 9	Los contenidos desarrollados en el Curso son realmente específicos y no los había estudiado en la carrera.
DC 10-17	Objetivos Específicos.
DC 18	Creo que, en general, la calidad del Curso responde a las expectativas que yo tenía cuando me matriculé en él.
DC 19	La formación recibida me ha proporcionado medios para la mejora de mi trabajo actual y/o futuro.
DC 20	La calidad del Curso se corresponde con el precio de la matrícula.
DC 21	¿Crees que se podría mejorar algunas cosas de este Curso? Indícalas.

4. RESULTADOS

Teniendo en cuenta el espacio disponible para la presentación de este trabajo, se presentan los resultados de los análisis realizados tomando como «organizador» los análisis canónicos (a) y especificando los resultados de los análisis univariados (b) a propósito de las relaciones descritas en cada dimensión.

Indicadores absolutos

Se extraen dos dimensiones canónicas significativas de acuerdo con la prueba de Bartlett —ver tabla 2 izquierda—. En la *Primera Dimensión* se asocian todas las variables de ambas series excepto Coste. De su composición se desprende que los cursos en los que se identifica un mejor Logro de Objetivos y una calidad global mayor son aquéllos que se estructuran con un menor número de materias, especialmente, junto a un menor número de profesores, materias y horas; es decir, los programas de tamaño más reducido, que corresponden a programas más especializados, son los que obtienen mejores valoraciones de sus estudiantes.

La síntesis de relaciones identificada en esta dimensión se traduce en términos específicos, teniendo en cuenta el comportamiento concreto en los ítems detectado en los *estudios univariados* —ver tabla 3— en los siguientes elementos:

a) Los estudiantes de *cursos con menor número de materias* valoran más positivamente el cumplimiento con el Calendario, Horarios y Cursos previstos ($\alpha=0.001$). A su vez, valoran también más positivamente la adecuación de los materiales disponibles en las aulas, así como son los que menor número de aspectos señalan que hay que mejorar ($\alpha=0.01$). Por otra parte, también son éstos los que valoran más positivamente la adecuación de los materiales de trabajo ofrecidos por el profesorado ($\alpha=0.001$). Finalmente, presentan una valoración más positiva en todos los ítems relacionados con la Utilidad global del curso ($\alpha=0.001$) y el Logro de Objetivos.

b) Los estudiantes que asisten a *cursos de menor duración* (Indicador N° de Horas) valoran más positivamente la adecuación de la infraestructura de aulas y del material disponible en las mismas ($\alpha=0.001$), así como ofrecen valoraciones más elevadas en todos los ítems relacionados con la adecuación de la actuación del Profesorado, Logro de Objetivos y Utilidad global del curso —ver niveles de significación en tabla 3—.

c) Por último, los estudiantes asistentes a *cursos impartidos por menor número de profesores*, valoran más positivamente diversos aspectos tanto referidos a la organización general del curso como a su desarrollo. Así, respecto a Aspectos Organizacionales, opinan más favorablemente en cuanto a la información facilitada respecto a la organización general del curso ($\alpha=0.05$), la adecuación de la infraestructura de aulas ($\alpha=0.01$), así como son los que menor número de recomendaciones de mejora en este aspecto realizan ($\alpha=0.01$). Por otra parte, opinan más favorablemente en cuanto a la claridad en el planteamiento de objetivos por parte del profesorado y su coordinación ($\alpha=0.001$), el Logro de Objetivos ($\alpha=0.001$) y la utilidad de los contenidos del curso para el ejercicio profesional ($\alpha=0.05$).

Respecto a la *Segunda Dimensión*, en ella se asocian los siguientes componentes de evaluación e indicadores: la adecuación de la actuación del Profesorado muestra una relación directa con el Número de Alumnos, al tiempo que una relación inversa con el Coste, particularmente, así como con el Número de Profesores y de Horas; es decir, los cursos de menor coste y mayor demanda, más cortos y con menor número de profesores son los que presentan mayor satisfacción con la actuación del profesorado.

En *términos univariados*, la relación descrita entre Indicadores y Componentes de Evaluación, se especifica en relación al comportamiento observado en los ítems en los siguientes aspectos:

a) La asociación más significativa se produce entre la valoración de la adecuación del Profesorado y Coste. Así, en los *cursos de menor coste* se produce una valoración significativamente mejor ($\alpha=0.001$) en el 80% de los ítems relativos a la adecuación del profesorado.

b) En los *cursos con menor número de profesores* también se produce una mejor valoración del profesorado, en concreto en cuanto a su claridad en la exposición de objetivos y en su coordinación ($\alpha=0.001$).

c) De igual modo, en los *cursos de menor número de horas totales* se valora mejor al profesorado, observándose diferencias significativas en todos los ítems considerados ($\alpha=0.001$).

d) Por último, también se observa una mejor valoración del profesorado en los *cursos de mayor matrícula* , habiéndose encontrado diferencias significativas al nivel del 0.001 en el 60% de los ítems y al 0.05 en un ítem.

TABLA 2
 DIMENSIONES CANÓNICAS EXTRAÍDAS EN AMBOS ESTUDIOS (IZQUIERDA INDICADORES ABSOLUTOS/DERECHA INDICADORES DE RAZÓN)

	1	2		1	2	3
1ª SERIE			1ª SERIE			
Organización	0,54	0,21	Organización	0,60	0,68	0,40
Profesor	0,65	0,73	Profesor	0,22	0,82	-0,51
Utilidad y Contenidos	0,66	-0,05	Utilidad y Contenidos	0,83	0,23	-0,40
Logro Obj.	0,94	-0,30	Logro Obj.	0,76	-0,04	-0,43
2ª SERIE			2ª SERIE			
Nº Horas	-0,38	-0,44	Horas/Mater.	0,91	0,04	-0,12
Nº Prof.	-0,48	-0,69	Horas/Prof.	0,03	-0,37	-0,80
Nº Alumnos	0,33	0,58	Materias/Prof.	-0,79	0,23	-0,36
Coste	-0,01	-0,81	Coste/Alum.	0,23	-0,89	0,16
Nº Materias	-0,91	0,25	Coste/Horas	0,30	-0,69	0,14

Indicadores de Ratio

Se han identificado tres dimensiones significativas de acuerdo con la prueba de Bartlett —ver tabla 2 derecha—. La *Primera Dimensión* es de carácter general y en ella se asocian, con relación directa, la satisfacción de los estudiantes respecto a la Utilidad global del curso, Logro de Objetivos y adecuación de la Organización, con la Ratio de Horas por Materia. Con relación inversa se asocia la Ratio Materias por Profesores. De este modo, los cursos globalmente mejor valorados, excepto en lo referente a la actuación del profesorado, son aquéllos que disponen de materias más extensas en su duración y que, a su vez, están impartidas por menor número de profesores.

Profundizando en el carácter de estas relaciones a partir de los *contrastes realizados por ítems* , pueden observarse las siguientes notas características:

a) La relación entre Utilidad global del curso y Logro de Objetivos con las horas destinadas a cada materias, se concreta en la observación de diferencias significativas en cuanto a la especificidad de los contenidos impartidos en el curso ($\alpha=0.001$) y a su utilidad para la mejora profesional ($\alpha=0.05$), a favor de los *cursos con materias de duración superior a 4 créditos* .

b) También este grupo presenta una mejor valoración del cumplimiento del calendario previsto y de la suficiencia de tiempo destinado a cada materia ($\alpha=0.05$). Sin embargo, en los cursos que contienen materias de tamaño inferior a 4 créditos, se han valorado más positivamente los aspectos relativos a la infraestructura (ítems 3 $\alpha=0.001$ — y 4 $\alpha=0.01$).

c) Por último, los tres componentes evaluativos que se integran en la dimensión (Utilidad, Logro de Objetivos y Organización) presentan, en términos generales, mejores niveles de valoración en aquellos *cursos en los que cada materia la imparten menos de 2 profesores* . Así, se han observado

diferencias significativas en prácticamente todos los ítems referidos a estos componentes —ver tabla 4—.

La *Segunda Dimensión* incluye la adecuación del Profesorado y la Organización General del Curso, con relación inversa a los Costes por Alumno y Hora; es decir, se ha valorado más positivamente al profesorado y la organización de los cursos más baratos.

En cuanto al estudio univariado relativo a los ítems de adecuación del profesorado y organización respecto a los Indicadores de Coste, se han observado las siguientes diferencias:

a) Se produce una mejor valoración en todos los ítems relativos al profesorado ($\alpha=0.001$) en los cursos de menor coste por alumno, y respecto a la indicación de problemáticas docentes ($\alpha=0.001$) en los cursos de menor coste por hora.

b) En cuanto a aspectos organizacionales en los cursos de menor coste por alumno se observa una mejor valoración de la información facilitada acerca del curso y la organización general, así como se realiza menor número de recomendaciones de mejora ($\alpha=0.001$).

c) Sin embargo, en cuanto al Indicador de Coste por Hora, aunque a nivel global se asocia en la dimensión, no se observan más diferencias significativas.

TABLA 3

SIGNIFICACIÓN ALCANZADA EN FUNCIÓN DE LOS INDICADORES EN EL CUESTIONARIO DE ASPECTOS DE ORGANIZACIÓN Y DESARROLLO DEL CURSO CON LA PRUEBA U DE MANN-WHITNEY

INDICADORES ÍTEMES	Nº. HORAS		Nº. PROFESORES		Nº. ALUMNOS		COSTE		Nº. MATERIAS	
	SIGNIFICACIÓN		SIGNIFICACIÓN		SIGNIFICACIÓN		SIGNIFICACIÓN		SIGNIFICACIÓN	
	Mayor	Menor	Mayor	Menor	Mayor	Menor	Mayor	Menor	Mayor	Menor
AO1				*		***				***
AO2						***				***
AO3		***		**						
AO4		***								**
AO5						***				***
AO6						**			**	***
AO7										
AO8		***		**		**			***	**
DC1		***		***		***			***	
DC2		***		***		***			***	
DC3		***				***			***	
DC4		***								
DC5		***				*			***	
DC6						***				***
DC7		***				***			***	
DC8		***		*		*				*
DC9	***		***			***				
DC10-17		***		***		***				***
DC18		*				***				***
DC19		***				*				***
DC20			***			**				***
DC21		**								***

En la *Tercera Dimensión*, las correlaciones canónicas de las variables originales de la primera serie (componentes evaluativos) son más bajas en términos generales que en las restantes dimensiones; sin embargo, estimamos que presentan tendencias claras en cuanto a su contenido. Así, se asocian en el mismo polo de la dimensión, la adecuación del Profesorado, el Logro de Objetivos y la Utilidad global del curso, con la Ratio Horas por Profesor. En sentido inverso, se asocia la valoración de la organización general del curso. En este sentido, puede entenderse que en los cursos en que se producen relaciones más extensas de los profesores con los estudiantes se aprecia en éstos una mejor opinión de los factores descritos.

Respecto a los contrastes univariados realizados con los ítems de adecuación del profesorado, logro de objetivos y utilidad global en relación al indicador Ratio Horas/Profesor, se han obtenido los siguientes resultados:

a) Los estudiantes de cursos en los que los profesores, por término medio, imparten más de 17 horas valoran más positivamente la coordinación entre los profesores ($\alpha=0.001$), si bien en los cursos en que

TABLA 4
SIGNIFICACIÓN ALCANZADA EN FUNCIÓN DE LOS INDICADORES DE RAZÓN EN EL CUESTIONARIO DE ASPECTOS DE ORGANIZACIÓN Y DESARROLLO DEL CURSO CON LA PRUEBA U DE MANN-WHITNEY

INDICADORES ÍTEMS	Nº. HORAS		Nº. PROFESORES		Nº. ALUMNOS		COSTE		Nº. MATERIAS	
	Mayor	Menor	Mayor	Menor	Mayor	Menor	Mayor	Menor	Mayor	Menor
AO1									***	
AO2			*			*			***	
AO3		***				***				
AO4		**		*		**				
AO5	*		*			**				
AO6						*	*			
AO7	*		***			***				
AO8									***	
DC1									***	
DC2			***			*			***	
DC3									***	
DC4						*				***
DC5						**			***	***
DC6						***				
DC7						*			***	
DC8						***				
DC9	***									
DC10-17						*				***
DC18						***				
DC19	*					***				
DC20				**		***				
DC21						***				

la dedicación media del profesor es menor se valora más positivamente la relación calidad-precio ($\alpha=0.01$).

b) En cuanto a Aspectos organizacionales si bien la tendencia global de asociación es clara, en cuanto a sus especificaciones por ítems presenta resultados contradictorios —ver tabla 4—, pudiéndose entender su participación en la dimensión de carácter residual.

5. CONCLUSIONES

Aunque, como es obvio, los indicadores estructurales considerados no pueden explicar por sí solos la calidad percibida, puede afirmarse que existe una cierta dependencia entre ambos elementos. En este sentido, estimamos de interés continuar explorando el rol de algunos de estos indicadores y sus relaciones con otros que aquí no se han considerado.

Atendiendo el valor diferencial de los indicadores analizados puede concluirse que globalmente los indicadores absolutos tienen mayor dependencia que los ratios. No obstante, existen conclusiones coincidentes para ambos conjuntos, como por ejemplo la estructura dimensional canónica básica, en la que se observan —con ambos conjuntos de indicadores— tendencia similares. Así, la dependencia de la valoración de la Utilidad global del Curso, el Logro de Objetivos y la Organización general del mismo del conjunto de indicadores —salvo Coste—, por un lado, y, por otro, la relación entre Coste/s y adecuación del Profesorado.

Respecto a esta segunda dimensión, hay que señalar que, en general, los Cursos de menor coste responden preferentemente a ofertas docentes soportadas directamente desde los Departamentos, con baja o nula participación de docentes de otras Universidades y/o extrauniversitarios y se ajusten en mayor medida a la organización derivada de la primera dimensión.

En cualquier caso, creemos necesario comprobar la permanencia del modelo de relaciones descrito a través de las Áreas, lo cual no se ha podido comprobar por limitaciones muestrales.

En cuanto a líneas organizativas, pueden extraerse las siguientes conclusiones:

a) Considerando que los indicadores absolutos sirven para caracterizar la tipología de oferta de los Cursos estudiados, los Cursos que reciben una mayor valoración, en términos generales, serían aquellos de menor duración y mayor especialización de contenido y, con una estructura organizativa más simple. La valoración del profesorado aunque relacionada globalmente con la valoración general del programa, se puede considerar, en parte, independiente de ésta y se relaciona fundamentalmente con los indicadores de coste; de forma que el perfil de Cursos en los que se da una mayor valoración del profesorado, son los de menor coste, con mayor número de alumnos, de menor duración y en los que participan menor número de profesores.

b) Por otra parte, teniendo en cuenta que los indicadores Ratio equiparan la oferta y describen en términos comparables los Cursos en cuanto a su estructura interna, se pueden extraer las siguientes consideraciones:

b.1) Las mejores valoraciones en cuanto a la Organización general, Logro de Objetivos y Utilidad global del Curso, se producen en aquellos Cursos que presentan una organización clásica, es decir, con materias de mayor duración impartidas fundamentalmente por un profesor.

b.2) Como en el caso de los indicadores absolutos, también se observa aquí la relación entre la valoración del profesorado y el Coste; de forma que resulta ésta mejor en los Cursos de menor Coste.

6. REFERENCIAS

CAVE M., HANNEY S. y KOGAN, M. (1991): The use of performance indicators in Higher Education (2ª Edición). Jessica Kingsley Publishers, London.

- DOCHY, F. J. R. C., SEGERS, M. S. R., WIJNEN, W. H. F. W. (1990): Selecting Performance Indicators. En GOEDEGEBUURE, L. C. J., MAASSEN, P. A. M. y WESTERHEIJDEN, D. F. (Eds.): Peer Review and performance indicators. Uitgeverij Lemma B.B., Utrecht.
- GOEDEGEBUURE, L. C. J., MAASSEN, P. A. M. y WESTERHEIJDEN, D. F. (Eds.) (1990): Peer Review and performance indicators. Uitgeverij Lemma B.B., Utrecht.
- JOHNES, J. y TAYLOR, J. (1990): Performance indicators in Higher Education. SRHE and Open University Press, Buckingham.
- MORA, J. G. (1991): La evaluación de las instituciones universitarias. *Revista de Investigación Educativa*, 9, 27-48.
- PÉREZ CARBONELL, A., JORNET, J. M.; SUÁREZ, J. M. y FERRÁNDEZ, M. R. (1993): Diseño de un sistema de evaluación para Cursos de Postgrado en la Universitat de València. Vicerrectorado de Estudios, Universitat de València.

FACTORES INTERVINIENTES EN LA VALORACIÓN DEL PROFESOR POR PARTE DEL ESTUDIANTE

J. González Such, J. M. Jornet, A. Pérez Carbonell y M. R. Ferrández
Universitat de València*

1. INTRODUCCIÓN

En la literatura sobre evaluación del profesorado universitario a partir de opiniones de estudiantes se recogen múltiples estudios que, en diferentes contextos, intentan relacionar los resultados de estas evaluaciones con diferentes variables, bien como factores de eficacia docente, bien como posibles fuentes de sesgo (De Neve y Jansen, 1982; Cranton y Smith, 1986; Marsh, 1987; Dickinson, 1990; Potter y Emanuel, 1990; Goldberg y Callahan, 1991; West y Denton, 1991, por citar algunos). Una lectura detenida de este tipo de estudios pone de manifiesto la inconsistencia de resultados para identificar los factores que influyen en la evaluación, así como induce a pensar en la enorme influencia de las características socio-educativas de los respondientes, de la cultura organizacional en que se inserta el sistema de evaluación, los usos particulares que se den a éste, etc. Todo ello implica que los factores que influyen en los resultados de la evaluación deben ser identificados para cada sistema particular y que, si bien la abundancia de resultados presente en la literatura evaluativa puede orientar acerca de los patrones generales, estos no pueden asumirse de forma inequívoca para todo sistema de evaluación.

2. OBJETIVOS

En este sentido, con el uso del programa de evaluación del profesorado universitario desarrollado por la Universitat de València desde el curso 1986-87, se ha ido observando una serie de inconsistencias en los patrones de respuesta a los cuestionarios de opinión por parte de los estudiantes, como por ejemplo, la presencia de sets de respuesta, la enorme variabilidad interna en la evaluación de un mismo profesor, con la presencia de evaluaciones totalmente contradictorias, etc.

El estudio que aquí se presenta forma parte de una programa más amplio, de carácter exploratorio, orientado a identificar hipótesis de trabajo en torno a factores relacionados con la evaluación a partir de opiniones de estudiantes, así como fuentes de sesgo en los mismos.

* Avda. Blasco Ibáñez, 21. 46010 VALENCIA. Teléfono (96) 3864420/30.

3. METODOLOGÍA

3.1. Procedimiento de recogida de información y muestra

Este estudio se basa en un programa de entrevista desarrollado en la Universitat de València estructurado en dos fases:

a) Primera fase: Fue desarrollada durante el curso 1990-91 y se basó en un modelo de entrevista semiestructurada con la que se trabajó con 200 estudiantes de diferentes facultades de la Universitat de València. La finalidad era identificar la existencia de constructos subyacentes a la evaluación del profesorado realizada a partir de cuestionarios de estudiantes. Como constructos se perseguía identificar conductas y/o comportamientos docentes que fueran especialmente relevantes para definir la opinión del estudiante respecto de sus profesores. En el proceso de elicitación se utilizaron como Elementos ítems del Cuestionario estándar de Evaluación del Profesorado a partir de opiniones de estudiantes. La estrategia básica en el proceso de entrevista incluyó los siguientes pasos: a) Evaluar al mejor profesor que tuvo en el curso, b) evaluar al peor profesor c) Reflexionar para cada ítem en qué diferían ambos profesores, qué avalaba las diferencias de puntuación que habían otorgado. Como resultado de estas entrevistas se realizó un listado de los 100 constructos más frecuentes, el cual fue reducido, tras sucesivas revisiones por un grupo de expertos —teniendo en cuenta su afinidad— a un total de 56 constructos. Como característica central de los mismos se pretendió que su definición fuera de carácter bipolar y que, en lo posible, ambos polos tuvieran la misma carga, es decir, no tuvieran ningún tipo de connotaciones positivas o negativas.

b) Segunda fase: Se desarrolló durante el curso 1991-92 y se basó en un modelo de entrevista estructurada construida con dos partes diferenciadas: una inicial, en la que los estudiantes debían verbalizar las conductas del profesorado que influían en la puntuación que otorgaban al profesor en cada uno de los ítems del Cuestionario de Evaluación del Profesorado por parte de los estudiantes. La segunda parte se construyó a partir de las entrevistas efectuadas en la primera fase ya citada y quedó configurada por los diferentes constructos elicitados por los estudiantes. Esta entrevista fue aplicada a 125 estudiantes elegidos aleatoriamente de la Universitat de València, de los cuales fueron eliminados 21 quedando finalmente 104.

3.2. Análisis

En este trabajo se reseñan los resultados relativos al estudio descriptivo de las respuestas sobre los constructos seleccionados. Asimismo, se aporta información acerca de la consistencia del patrón global, analizado mediante análisis Cluster¹. Por otra parte se han considerado tres variables de estratificación del grupo de estudio: a) Área (Experimentales, Salud, Sociales, Humanidades y Psicopedagógicas), b) Sexo, c) Ciclo (1º, 2º). Entre los grupos formados a partir de cada una de estas variables se ha analizado la existencia de diferencias significativas mediante análisis de varianza.

1 Se ha utilizado el programa QUICK-CLUSTER del paquete SPSS 5.0.

4. RESULTADOS

Dos cuestiones previas

Al ser estos resultados una parte del programa de entrevista realizado estimamos conveniente aportar previamente información acerca de dos aspectos que estimamos de interés con objeto de contextualizar más adecuadamente los restantes resultados.

En primer lugar, la totalidad de estudiantes entrevistados señalaron —en cuestiones abiertas— que los aspectos —o dimensiones clave— en las que se basan para evaluar al profesorado son su estilo docente (forma/s de dar clase, de atender a alumnos, etc.) y la relación que establece con los estudiantes y el conocimiento de la materia.

En segundo lugar, una inconsistencia observada reiteradamente en las evaluaciones la constituye el hecho de encontrar puntuaciones negativas en cuestiones de cumplimiento de obligaciones (asistencia a clase, puntualidad, cumplimiento del horario de atención a estudiantes) en casos que no se habían producido incumplimientos. En este sentido ante la cuestión de si habían evaluado de forma diferente a lo real en estos ítems en alguna ocasión, entre el 10-15% —según áreas— de los estudiantes consultados señalaban haberlo hecho en alguna ocasión; y, la misma cuestión referida a si conocen compañeros que lo hayan hecho, alcanza niveles de hasta un 25% —según áreas—. En ambos casos, el sentido del «cambio» —recogido en cuestiones abiertas— ha sido negativo y, los motivos giran en torno al uso de la «evaluación como castigo». En este sentido, aunque estos resultados son aún provisionales, hay que señalar que si bien este hecho puede considerarse una importante fuente de sesgo parece afectar preferentemente a la fiabilidad y validez de las evaluaciones negativas.

Características de los constructos

Un primer elemento denotativo de este acercamiento lo constituyen los constructos mismos, su generalidad. En este sentido, hay que señalar que los estudiantes entrevistados, tanto en la primera como en la segunda fase, tendían a identificar como constructos los elementos mismos, es decir, solían definir las razones de su evaluación a partir de los mismos ítems, siendo poco frecuente que aportaran información adicional más concreta, relativa a conductas docentes o motivaciones propias.

Valoración de los constructos

En las Tablas 1 a 5 se recogen los constructos utilizados en la segunda fase. El perfil viene determinado por la media estimada para el grupo total. Aunque para esta presentación se ha utilizado una escala de 5 puntos, la cuantificación inicial en las entrevistas se realizó con sentido bipolar, con un punto medio en 0 —«No influye cunado evaluó»— y los extremos en 2 —«Valoro siempre mejor»—.

En el primer grupo, *características de las asignaturas*, los estudiantes señalan que tienden a valorar mejor casi siempre a los profesores cuando imparten asignaturas que les gustan y/o que son importantes para el posterior desarrollo profesional o lo son en el conjunto de la carrera (Medias= 2.17, 2.31 y 2.37, respectivamente). Por otra parte, señalan que en su evaluación del profesor no les influye la dificultad de la asignatura, el carácter de la misma (obligatoria/optativa) o el hecho de si está o no masificada (Medias= 3.28, 3.06 y 3.35, respectivamente). Todas estas opiniones se pueden valorar como muy homogéneas (con Cocientes de variación entre 20.26% y 37.39%) y son representativas para el conjunto de sujetos entrevistados, no habiéndose encontrado diferencias estadísticamente significativas en ninguno de los contrastes realizados entre áreas, ciclos o sexo —ver tabla 1—.

El segundo grupo, formado por los constructos referentes al *Estilo Docente* del Profesor según los

TABLA 1
FACTORES REFERIDOS A LA ASIGNATURA

	1	2	3	4	5	
La asignatura me gusta						La asignatura no me gusta
La materia es difícil						La materia es fácil
El número de alumnos es elevado						El número de alumnos es escaso
La asignatura es importante para el posterior desarrollo profesional						La asignatura es accesorio para el posterior desarrollo profesional
La asignatura es importante en el conjunto de la carrera						La asignatura es accesorio en el conjunto de la carrera
La asignatura es obligatoria						La asignatura es optativa

TABLA 2
FACTORES REFERIDOS AL ESTILO DOCENTE

	1	2	3	4	5	
En sus explicaciones siempre se ciñe a los contenidos de la asignatura						En sus explicaciones se extiende hacia otros temas, aunque sean de otras asignaturas
La actividad docente de esta asignatura va dirigida únicamente a aprobar los exámenes						La actividad docente de la asignatura va dirigida a una formación global
La asistencia a clase me resulta útil para adquirir las ideas clave						Las ideas clave se pueden entender aunque no se asista a clase
La asistencia a clase me resulta útil para aclarar puntos oscuros						La asistencia a clase es poco importante para ir bien en la asignatura
Utiliza un lenguaje sencillo y coloquial						Utiliza un lenguaje técnico especializado
El profesor pregunta si está claro lo que explica y da opción a preguntar dudas						El profesor sigue un ritmo constante de explicación en el que no cabe interrumpir para preguntar dudas
El profesor responde siempre, aunque su respuesta no sea la correcta						El profesor reconoce su falta de información sobre una duda
Contesta con rapidez las preguntas que realizamos						Tarda en contestar las preguntas que se le formular.
En sus explicaciones se esfuerza por adaptarse al nivel del alumno						En sus explicaciones aspira a que nos ajustemos a su nivel
Le entendemos perfectamente en sus explicaciones, es un buen orador, se explica con claridad						Le entendemos perfectamente en sus explicaciones, aunque no es un buen orador, se explica con claridad
Explica todo lo necesario aunque tenga que clarificar lo de niveles inferiores						Explica lo que corresponde a su nivel
En sus explicaciones se ajusta al temario de la asignatura						Explica sólo algunos temas de los que componen el temario
Explica fundamentalmente algunos temas y conceptos						Explica todos los conceptos por igual
El ritmo de clase permite tomar apuntes casi literalmente						El ritmo de clase no permite tomar apuntes literales
El profesor nos comenta las últimas invest. técnicas y trabajos sobre la materia						El profesor se limita a explicar los contenidos de la materia
Utiliza distintos tipos de clase						Su forma de dar la clase es siempre la misma, no introduce cambios
Sus clases parecen estar muy preparadas						Parece que sus clases se basan en su capacidad de improvisación
Reparte el trabajo de clase entre sus intervenciones y las de los estudiantes						El tiempo de clase se llena con la explicación que realiza el profesor
El profesor utiliza los mismos apuntes que los de años anteriores						El profesor cambia cada año -todo o parte- sus apuntes
El profesor utiliza medios audiovisuales						El profesor no utiliza medios audiovisuales
El profesor tiene en cuenta las opiniones, sugerencias o intereses de los alumnos para estructurar la clase						El profesor tiene su propio esquema de clase y no pide opiniones o sugerencias de los alumnos
El profesor integra información sobre salidas profesionales o temas relacionados con la docencia de la asignatura						El profesor no integra información sobre salidas profesionales o temas relacionados con la docencia de la asignatura
Se puede hablar con el profesor fuera del aula y del horario de atención a alumnos						Sólo se puede hablar con el profesor dentro de los horarios establecidos

TABLA 3
FACTORES REFERIDOS A MATERIALES Y PROGRAMA

	1	2	3	4	5	
La bibliografía está accesible en Biblioteca o librerías normales		•				La bibliografía no está en Bibliotecas o librerías normales
El profesor da a conocer, comenta el programa de la asignatura		•				El profesor no comenta el programa
La asignatura tiene programa		•				La asignatura no tiene programa
El programa de la asignatura incluye criterios de evaluación		•				El programa de la asignatura no incluye criterios de evaluación
El programa de la asignatura incluye temario		•				El programa de la asignatura no incluye temario
El programa de la asignatura incluye bibliografía		•				El programa de la asignatura no incluye bibliografía

TABLA 4
FACTORES REFERIDOS A LA EVALUACIÓN

	1	2	3	4	5	
El profesor exige poco			•			El profesor exige mucho
Los contenidos de examen son siempre temas importantes		•				Suele sorprender incluyendo en el examen temas poco o nada importantes
La calificación incluye además del examen otras posibilidades, como trabajos, seminarios, etc.		•				La calificación se obtiene únicamente por el examen
El examen es de desarrollo			•			El examen es tipo test
El examen es fundamentalmente teórico			•			El examen es fundamentalmente práctico
El profesor accede a la revisión de las calificaciones si considera que hay error	•					El profesor no accede fácilmente a la revisión de las calificaciones
Los contenidos de examen se ajustan únicamente a lo dado en clase		•				Los contenidos de examen incluyen otros conceptos no dados en clase
El profesor entrega los resultados de los exámenes en un tiempo corto		•				El profesor tarda en entregar los resultados de los exámenes
Los resultados del examen influyen en la forma en que el profesor enfoca el curso			•			Los resultados de los exámenes no influyen en el desarrollo posterior de la asignatura
Los exámenes son largos			•			Los exámenes son cortos
El profesor ofrece un periodo para comentar la valoración de los exámenes		•				El profesor no ofrece un periodo para comentar la valoración de los exámenes

entrevistados puntúan siempre mejor a los profesores que «preguntan si está claro lo que explica y da opción a preguntar dudas» (Media= 1.39), los que «se esfuerzan por adaptarse al nivel del alumno» (Media= 1.5). Los estudiantes han opinado que puntúan mejor casi siempre al profesor que en sus explicaciones se ciñe a los contenidos de la asignatura (Media= 2.3), cuya actividad docente va dirigida a una formación global (Media= 3.88), siendo además este constructo significativo tanto a nivel de Ciclo incrementándose la tendencia en el segundo ciclo (0.01), como de Áreas (0.05), con medias que oscilan entre 3.4 del área de Ciencias Sociales hasta 4.28 del área de Psicopedagógicas. Los alumnos también consideran que puntúan mejor casi siempre a los profesores de las asignaturas en las que la asistencia a clase es importante para adquirir ideas clave (Media= 2.03) o para aclarar puntos oscuros (Media= 1.83), reconocen su falta de información sobre una duda (Media= 4.08), o contesta con rapidez las preguntas que realizan (Media= 2.37). También aparecen como mejor puntuados casi siempre los profesores que utilizan un lenguaje sencillo y coloquial (Media= 1.81), los que se les entiende perfectamente sus explicaciones, son buenos oradores, se explican con claridad (Media= 1.72), así como los profesores que explican lo necesario aunque tengan que clarificar lo de niveles inferiores (Media= 1.59), los que se ajustan al temario de la asignatura (Media= 1.89, significativa por ciclos al 0.05, siendo los alumnos de primer ciclo los que mayor puntuación otorgan al profesor que se

TABLA 5
FACTORES REFERIDOS A LAS CARACTERÍSTICAS DEL PROFESOR

	1	2	3	4	5	
El profesor tiene un carácter cordial, abierto, establece fácilmente una relación personal con los alumnos		•				El profesor que por su carácter serio, mantiene una relación estrictamente profesional con sus alumnos
El profesor parece importante; su prestigio profesional va más allá de su actividad docente			•			El profesor se ocupa fundamentalmente de su docencia. Es poco conocido fuera de la Universidad.
El profesor que parece que tiene más experiencia docente			•			El profesor que parece que tiene menos experiencia docente
El profesor más viejo			•			El profesor más joven
El profesor es de mi propio sexo			•			El profesor de un sexo diferente al mío
El profesor entiende las lenguas de la Comunidad			•			El profesor entiende únicamente una lengua
Políticamente, parece conservador			•			Políticamente, parece progresista
Parece que sus ideas políticas son de izquierda			•			Parece que sus ideas políticas son de derecha
Es muy crítico con el sistema establecido			•			Está de acuerdo con el sistema establecido
Parece muy racional			•			Parece muy emocional

ajusta al temario), explica todos los conceptos por igual (Media= 3.97), siguen un ritmo de clase que les permite tomar apuntes literales (Media= 1.59) y los que comentan las últimas investigaciones técnicas y trabajos sobre la materia (Media= 1.67, significativo por Ciclos al 0.05, siendo los de segundo ciclo los más cercanos a este polo). De igual forma, encontramos que los profesores que son puntuados mejor casi siempre son los que utilizan distintos tipos de clase (Media= 1.52), los que se preparan las clases (Media= 1.95), motivan a sus alumnos para que participen en sus clases (Media= 1.81), los que renuevan cada año sus apuntes (Media= 3.58), utilizan medios audiovisuales (Media= 2.28), tienen en cuenta las opiniones, sugerencias o intereses de los alumnos para estructurar la clase (Media= 1.58), los que integran información de salidas profesionales o temas relacionados (Media= 1.56) y los que sus alumnos pueden hablar con ellos fuera del aula y del horario de atención a alumnos (Media= 1.67) —ver tabla 2—.

Considerando el tercer grupo, factores referentes a *materiales y programa*, encontramos que los estudiantes puntúan mejor casi siempre a los profesores que tienen programa de la asignatura (Media= 1.83), lo dan a conocer (Media= 1.78) y este programa incluye: temario (Media= 1.94, con diferencias significativas a un nivel de 0.05 por ciclos, siendo los de primer ciclo los que más se acercan a este polo), criterios de evaluación (Media= 2.02), bibliografía (Media= 1.86) y está en librerías usuales (Media= 1.72).

Considerando el grupo de factores referidos a la *evaluación*, los alumnos entrevistados puntúan más alto siempre a los profesores que acceden a la revisión de las calificaciones si considera que hay error (Media= 1.32). De igual forma, han puntuado que otorgan mejor puntuación casi siempre a los profesores en los que los contenidos del examen son siempre temas importantes (Media= 1.84), la calificación de los exámenes incluye otras posibilidades como trabajos, seminarios, etc. (Media= 2.03, significativa a un nivel del 0.01 por Ciclos y siendo los de segundo ciclo los que la consideran más importante), los contenidos de examen se ajustan a lo dado en clase (Media= 1.73), el profesor que entrega los resultados de examen en un tiempo corto (Media= 2.17, significativa a un nivel de confianza del 0.05 por áreas), que los resultados del examen influyen en la forma en que el profesor enfoca el curso (Media= 2.13) y aquellos que los alumnos tienen oportunidad de comentar la valoración de los exámenes (Media= 1.87).

Sin embargo, no tienen en cuenta los siguientes factores para valorar al profesor: si el profesor aprueba fácilmente o exige mucho (Media= 2.61, significativa a un nivel de confianza de 0.01 por

Ciclos, tendiendo el primer ciclo hacia el polo de facilidad). También opinan que no tienen en cuenta si el examen es de desarrollo o de tipo test, aunque con una ligera tendencia hacia el de desarrollo (Media= 2.59); si el examen es teórico o práctico (Media= 3, diferencias significativas a un nivel de conf. de 0.05 por áreas, siendo el área de Experimentales la que mayor importancia otorga al examen práctico y los alumnos de Sociales y de Salud los que más se decantan hacia un examen de tipo teórico).

Finalmente, respecto a los factores referidos a las *características del profesor*, los estudiantes puntúan casi siempre mejor a los profesores que tienen un carácter cordial, abierto, que establecen fácilmente una relación personal con los alumnos (Media= 1.55), así como al profesor que parece que tiene más experiencia como docente (Media= 2.49) y al profesor que entiende las lenguas de la Comunidad (Media= 2.3). Tampoco tienen en cuenta si el profesor parece racional o emocional (Media= 2.6), aunque en este factor hay diferencias significativas a un nivel del 0.05 por ciclos, puntuando mejor los estudiantes de segundo ciclo al profesor racional.

Análisis Cluster

La exploración mediante análisis Cluster revela la conformación de dos grupos claramente diferenciados (grupo I: 39 casos; grupo II: 65 casos). Entre ellos se constatan diferencias significativas en el 75% de los constructos considerados. Las características de los patrones de respuesta pueden sintetizarse en los siguientes puntos:

a) No corresponden a características de asignación social de los grupos: no hay diferencias significativas en las variables de clasificación (sexo, ciclo, área) entre ambos grupos.

b) Ambos patrones responden a «intensidad» de respuesta, siendo ésta sistemática y constante; de forma que puede entenderse que el segundo grupo presenta respuestas más extremas, más claras, en cuanto a identificación de los factores intervinientes en la evaluación.

c) Asimismo, este patrón sistemático en ambos grupos mantiene en todos los constructos las tendencias descritas como perfil general.

d) La puntuación del patrón global identificado también en la tipología confirma la generalidad del mismo.

5. CONCLUSIONES

El valor exploratorio de este estudio reside fundamentalmente en la posibilidad de derivar hipótesis de trabajo. En este sentido, las conclusiones que se presentan pueden entenderse como líneas de investigación subsiguiente.

En términos generales, la utilización de la evaluación como elemento de «castigo» al docente entendemos que desvirtúa el procedimiento y lo limita en su utilidad. En cualquier caso, la identificación de este tipo de sesgo puede comprobarse si se mantienen en los cuestionarios los ítems de Cumplimiento de Obligaciones, dado que son claramente objetivables, en el sentido que señalan Muñoz, García y Virgos (1991). No obstante, en casos de profesores problemáticos, si se han generado los sets de respuesta negativa, el uso formativo queda limitado en tanto que a partir de estos cuestionarios no se identifica el tipo de problemática —queda «enmascarada»— y el uso sumativo necesariamente queda invalidado —al faltar a la necesaria validez de todo proceso evaluativo—.

Respecto a los factores que influyen en las evaluaciones de estudiantes, teniendo en cuenta los constructos analizados se puede afirmar que estos tienden a valorar a sus profesores de forma global. El patrón de preferencias extraído a partir de los sujetos entrevistados presenta escasas diferencias significativas entre los grupos formados por las variables consideradas (área, ciclo y sexo). Este hecho avalaría el uso de un único cuestionario de amplio espectro utilizable para el conjunto de una Univer-

sidad (hecho además concurrente con otros estudios realizados en este mismo contexto —Jornet, Suárez y González Such, 1989).

El perfil de «preferencias» identificado responde básicamente a la docencia como relación interpersonal, es decir, el estilo docente que, según los estudiantes, tienden a valorar mejor es el que se basa en el respeto del profesor hacia sus estudiantes, considera sus necesidades para un adecuado aprendizaje, establece una buena relación personal, es buen comunicador y conoce su/s materia/s.

En este sentido, el patrón educativo entendemos que alude, de forma no sistemática, a factores de eficacia docente en lo que habrá que profundizar.

No obstante, en dicho patrón también se identifican posibles fuentes de sesgo (como el gusto del estudiante por la materia, la importancia de ésta en la carrera y/o en la profesionalización, el carácter del profesor, su experiencia...) que deberán ser adecuadamente analizados y que son concurrentes con otros estudios y revisiones realizadas (Marsh, 1987; Shepard y Trank, 1989; De la Orden, 1990; Salvador, 1990; Scherr y Scherr, 1990; Shapiro, 1990; Escudero, 1991, entre otros).

Teniendo en cuenta los diferentes aspectos comentados, estimamos que el uso adecuado de este tipo de cuestionarios de evaluación debe circunscribirse al de elemento que operativiza la detección de casos problemáticos, no de las características del problema y por ello se debe ser cauteloso en su uso como instrumento de apoyo a decisiones.

6. REFERENCIAS

- CRANTON, P. A. y SMITH, R. A. (1986): A new look at the effect of course characteristics on student ratings of instruction. *American Educational Research Journal*, 23, 117-28.
- DE NEVE, H. M. F., & JANSSEN, P. J. (1982): Validity of Student Evaluation of Instruction. *Higher Education*, 11(5), 543-52.
- DICKINSON, D. J. (1990): The relationship between ratings of teacher performance and student learning. *Contemporary Educational Psychology*, 15, 142-51.
- ESCUDERO, T. (1991): Enfoques modélicos en la evaluación de la enseñanza universitaria. Ponencia en las *III Jornadas Nacionales de Didáctica Universitaria «Evaluación y Desarrollo Profesional»*. Las Palmas de Gran Canaria, 23-26 de septiembre.
- GOLDBERG, G. y CALLAHAN, J. (1991): Objectivity of student evaluations of instructors. *Journal of Education for Business*, 66, 377-78.
- JORNET, J. M., SUÁREZ, J. M. y GONZÁLEZ SUCH, J. (1989): Estudio de validación de un cuestionario de valoración de la docencia universitaria por los estudiantes en un conjunto homogéneo de centros de la Universitat de València.
- MARSH, H. W. (1987): Student Evaluation of Teaching, en DUNKIN, M. J. (ed.): *The International Encyclopedia of Teaching and Teacher Education*, Pergamon Press, Oxford.
- MUÑIZ, J., GARCÍA, A. y VIRGOS, J. M. (1991): Escala de la Universidad de Oviedo para la evaluación del profesorado. *Psicothema*, 3, 269-281.
- ORDEN, A. de la (1990): Evaluación, selección y promoción del profesor universitario. *Revista Complutense de Educación*, 1, 11-29.
- POTTER, W. J. y EMANUEL, R. (1990): Student's preferences for communication styles and their relationship to achievement. *Communication Education*, 39, 3, 234-49.
- SALVADOR, L. (1990): *Los docentes universitarios exitosos desde la perspectiva del alumno: su caracterización psicopedagógica*. Tesis Doctoral no publicada. Universidad de Salamanca.
- SCHERR, F. C. y SCHERR, S. S. (1990): Bias in student evaluations of teacher effectiveness. *Journal of Education for Business*, 65, 356-58.

- SHAPIRO, E. G. (1990): Effect of instructor and class characteristics on students' class evaluations. *Research in Higher Education*. 31, 2, 135-48.
- SHEPHERD, G. J., & TRANK, D. M. (1989): Individual differences in consistency of evaluation: Student perceptions of teacher effectiveness. *Journal of Research and Development in Education*, 22(3), 45-52.
- WEST, S. S. y DENTON, J. J. (1991): An empirical validation of the instrument: Student perceptions of teaching effectiveness. Comunicación presentada al *Annual Meeting of the American Educational Research Association* (Chicago, IL., April 3-7).

SESGOS EN LA EVALUACIÓN DE LA INVESTIGACIÓN PEDAGÓGICA Y PSICOLÓGICA: UNA SUCINTA REVISIÓN INTEGRATIVA

Antonio Fernández Cano
Universidad de Granada

1. INTRODUCCIÓN

Mucho se ha escrito y discutido sobre el diseño de investigaciones en Pedagogía y Psicología, pero escasamente sobre la evaluación de las mismas. El estudio del tópico, bien a nivel de proyecto o de estudio ya realizado, es de alta relevancia ya que puede cumplir múltiples funciones; entre otras:

— Estilar los mejores proyectos e investigaciones para, en consecuencia, aprender de ellos y recompensarlos precedentemente.

— Detectar y evitar plagios, fraudes, errores y violaciones ético-deontológicas.

— Aceptar los manuscritos de más alta calidad para su posterior publicación y/o inserción meta-analítica.

— Marca una correcta productividad de centros, departamentos e individuos.

— Señalar frentes «calientes» de investigación posibilitando el planteamiento de agendas.

Diversas modalidades se han venido utilizando para evaluar investigaciones. Las dos más usuales son la **revisión por pares**, o evaluación inmediata-formal o a priori, y el **impacto de citación**, evaluación mediata-formal o a posteriori. Aquí nos vamos a centrar en la evaluación de la investigación a partir de la revisión por pares.

2. CONCEPTUALIZACIÓN

La política de poder de los «colegios invisibles» y los propios paradigmas deben ser respetados si se quiere obtener el reconocimiento de la calidad de una investigación. Múltiples son los factores de adecuación; algunos de ellos, totalmente ajenos a una estructura de poder justa (entiéndase justa como pública) y/o ajenos a la naturaleza del propio paradigma. Estos factores diferenciales de sesgo se ponen de manifiesto, sobre todo, en el proceso de revisión de pares, no ya como una escasa fiabilidad (acuerdo/concordancia bajo entre valoradores) sino como un prejuicio, bastante más fiable, contra o a favor de factores de nula validez anexos a la investigación.

El uso de cualquier otro criterio distinto del de calidad metodológica introduce pues una amenaza potencial a la validez de la investigación. En el fondo, la existencia de sesgos en la evaluación de la investigación obedece a la tenaz influencia de la metateorías justificacionales (basadas en la autoridad) tan comunes en las ciencias sociales.

Existen abundantes estudios, comparativos e incluso experimentales, que testimonian la existencia de sesgos conscientes o inconscientes en el proceso de revisión por pares; aunque la mayoría de las indagaciones sobre sesgos se han realizado con metodología correlacional, lo cual ha suscitado serios problemas de interpretación.

3. HALLAZGOS GENERALES

3.1. Sesgo relativo a la afiliación institucional del autor

Se detecta un sesgo de favoritismo para autores de instituciones de alto prestigio incluso en revisión ciega (¡); o sea, los autores pertenecientes a instituciones con alta reputación reciben revisiones más favorables.

Incluso, también se detecta, corroborando una posición teórica propia del relativismo cognitivo, que la predisposición de los revisores está más sesgada a favor de autores que proceden de instituciones similares a las suyas, a un mismo grupo, área o asociación ya comparten un mayor conjunto de creencias sobre lo que es una buena investigación. El sesgo aparece como un **efecto halo**, en el que la afiliación institucional del autor está altamente correlacionada con la posibilidad de publicar (recibir evaluación positiva).

El controvertido estudio de Peters y Ceci (1982) fue aleccionador. Sustituyeron nombre e instituciones de 12 artículos ya publicados y volvieron a remitirlos a las mismas prominentes revistas en que aparecieron. Ocho de ellos fueron rechazados en base a «serias deficiencias metodológicas».

Estamos ante una de las manifestaciones del famoso «**Efecto S. Mateo**»: un curioso patrón de interacción selección x afiliación institucional.

La alternativa clásica de usar revisión ciega, aunque cuestionable, parece ser la única solución.

3.2. Sesgo relativo al revisor

Existe la creencia de un sesgo supercrítico por parte del revisor por aparecer «brillante pero cruel». Tal sesgo se manifiesta en hechos tales como: aceptar estudios en base a creencias o teorías asumidas por el revisor, divergentes valoraciones diferenciales (académicos vs. aplicados, investigadores vs. decisores, teóricos vs. prácticos, solicitantes vs. no solicitantes,..). En definitiva, los juicios sobre los que evalúa el revisor varían según la teoría que abrace (Lord et al. 1979) y según su rol profesional (Hall et al. 1988; Vockell y Asher, 1974; Ward et al. 1975).

Tal sesgo podría minimizarse aumentando el número de revisores que valoran un mismo estudio y/o asignando aleatoriamente los estudios a un mismo conjunto de revisores.

3.3. Sesgo pro estudios positivos

También denominado sesgo del prejuicio contra la hipótesis nula y que se manifiesta, véase Greenwald (1975), o Sterling (1959), en:

— Los autores son escasamente proclives a remitir sus trabajos para publicación si no obtienen resultados estadísticamente positivos).

— Las revistas publican mayoritariamente (alrededor del 90%) estudios con resultados positivos.

El obtener significación estadística (resultados positivos) es casi un prerrequisito para recibir calificadas evaluaciones y, en consecuencia, publicar y/o conseguir aceptabilidad. Se obvia que el avance científico se logra, a menudo, más poderosamente, mediante el rechazo de teorías: y la

estrategia capital para conseguir ésto es demostrar que la relaciones predictas (hipótesis) no se verifican; lo cual requiere, por tanto, aceptar la hipótesis nula.

El obtener significación estadística se ha convertido en un erróneo «sine qua nom» de respetabilidad y aceptación científica.

Tal sesgo se ha tratado de evitar utilizando dos aproximaciones:

- Valorando los informes sin la sección de resultados y con la sección de discusión modificada,
- y
- Valorando sólo el diseño y la importancia del estudio.

3.4. Ausencia de estudios de réplica

Los estudios de réplica, detectados por los revisores o declarados por los autores, son escasamente publicados, desvalorados; cuando la replicación en ciencias físicas es habitual y necesaria. Sin embargo, las réplicas indirectas (algunas variaciones, a veces, mínimas y/o nimias) o no confesadas son práctica abrumadora en ciencias sociales hasta el punto de haber dado paso a modalidades de investigación de síntesis. (Véase, al respecto, Darr, 1987).

3.5. Relativos al sexo del autor/a y/o revisor/a

Ponderar si la evaluación podría estar sesgada en función del sexo del autor y/o del revisor ha sido una constante, en la valoración de la producción investigadora, desde el estudio seminal de Goldberg (1968). El hallazgo de este estudio manifestaba que «hay una tendencia entre las mujeres a subvalorar el trabajo de profesionales de su mismo sexo».

El trabajo de Goldberg (1968) ha sido objeto de varios estudios de réplica con resultados contradictorios. La tendencia, con el tiempo, parece haber sido valorar equilibradamente a mujeres-expertas frente a hombres-expertos, manteniéndose el sesgo entre hombres y mujeres, no expertos, a favor de ellos.

En los últimos estudios (Lloyd, 1990), los resultados evidencian una desconfirmación total del postulado de Goldberg en el sentido de que las revisoras aceptaron más manuscritos de autoras que de autores y que los manuscritos de mujeres fueron más aceptados por revisoras que por revisores. Sin embargo, el propio Goldberg ya insinuaba que tal sesgo era más función del área de investigación. El sesgo sería más fuerte en áreas dominadas por varones que en áreas neutras o dominadas por mujeres.

Para controlar tal sesgo, se instituyeron políticas no sexistas, revisión ciega, se alentó la participación de mujeres a través de, por ejemplo, el aumento de editoras en consejos de edición de revistas.

3.6. Al sexo del autor en políticas editoriales

Las tasas de publicación, según sexo, eran tiempo atrás ostensiblemente divergentes y favorecían a varones. Pero tal sesgo, no sólo se ponía de manifiesto en cuanto autoría, sino progresivamente en cada paso hacia la pirámide de la eminencia: primeros autores, citas y editores.

Desde el estudio de arranque de Teghtsoonian (1974), el porcentaje de mujeres ha ido progresivamente en incremento aunque las divergencias aún subsistan (Gannon et al. 1992).

4. CONCLUSIONES

La existencia de sesgos en la evaluación de la investigación a través de la revisión por pares parece evidente. Las sospechas de conservadurismo, endogamia y pesimismo en el proceso científico han

sido una crítica constante. Armstrong (1982) llegó a ironizar, al respecto, emitiendo su sarcástica «fórmula de autor» al objeto de que cualquier investigador obtuviese una valoración inmediata y positiva de su trabajo; a saber:

- * No considerar un problema importante.
- * No retar las creencias existentes.
- * No obtener resultados sorprendentes.
- * No usar métodos sencillos.
- * No revelar todo lo que se sabe o aconteció sobre/en el estudio.
- * No escribir con claridad.

Mucho habrá que estudiar e investigar al respecto en nuestro país. Parece irónico que el método científico se haya aplicado tan escasamente a la evaluación del producto de la actividad científica, no ya para seleccionar lo mediocre de lo apto; sino para seleccionar entre lo apto, lo mejor.

El hecho de conocer la existencia de presuntos sesgos quizá pueda ayudarnos a una más válida y fiable evaluación de nuestra labor investigadora.

BIBLIOGRAFÍA

- AMSTRONG, J. S. (1982): Research on scientific journal: Implications for editors and authors. *Journal of Forecasting*, 1, 83-104.
- DARR, R. (1987): Another look at Meehl, Lakatos, and the scientific practices of psychologist. *American Psychologist*, 42, 145-151.
- GANNON, L., LUCHETTA, T., RHODES, K., PARDIE, L. y SEGRIST, D. (1992): Sex bias in psychological research. Progress or complacency? *American Psychologist*, 47 (3), 389-395.
- GOLDBERG, P. A. (1968): Are some women prejudiced against women? *Transaction*, 5 (5), 28-30.
- GREENWALD, A. (1975): Consequences of prejudice against the null hypothesis. *Psychological Bulletin*, 81, 1-20.
- HALL, B. W., WARD, A. W. y COMER, C. B. (1988): Published educational research: An empirical study of its quality. *Journal of Educational Research*, 81 (3), 182-189.
- LLOYD, M. E. (1990): Gender factors in reviewer recommendations for manuscript publication. *Journal of Applied Behavior Analysis*, 23 (4), 539-543.
- LORD, C., ROSS, L. y LEPPER, M. (1979): Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*, 37, 2.098-2.109.
- PETERS, D. P. y CECI, S. J. (1982): Peer-review practices of psychological journals: The fate of published articles, submitted again. *The Behavioral and Brain Sciences*, 5 (2), 187-195, 255.
- STERLING, T. D. (1959): Publication decisions and their possible effects on inferences drawn from tests of significance - or vice versa. *Journal of American Statistical Association*, 54, 30-34. También en *The significance test controversy* (1970), D. Morrison y R. Henkel (Eds.), (pp. 295-300). Chicago.: Aldine.
- TEGHTSOONIAN, M. (1974): Distribution by sex of authors and editors of psychological journals, 1970-72: Are there enough women editors? *American Psychologist*, 29, 262-269.
- VOCKELL, E. L. y ASHER, W. (1974): Perceptions of document quality and use by educational decision makers and researchers. *American Educational Research Journal*, 11 (3), 249-258.
- WARD, A. W.; HALL, B. H. y SCHRAMM, C. F. (1975): Evaluation of published educational research: A national survey. *American Educational Research Journal*, 12 (2), 109-128.

INVESTIGACIÓN ASISTIDA POR ORDENADOR: ENFOQUE METODOLÓGICO

Guillem Bou Bauza y Montserrat Clotet Masana
Universidad Autónoma de Barcelona

En el *Laboratorio de Informática aplicada a la Educación*, hemos iniciado una línea de investigación recogiendo dos realidades ya asentadas: La **EAO** como intervención educativa y la aportación de experiencias en el campo de la recogida de datos para la investigación que se remontan a las realizadas por Norman en 1982. Bajo las siglas **IAO**, **Investigación Asistida por Ordenador**, presentamos dicho campo de estudio. Mostraremos qué es lo que nosotros entendemos por **IAO** exponiendo algunas de las investigaciones realizadas bajo esta metodología en nuestro laboratorio, hecho que nos permitirá observar cuales son las ventajas que ésta aporta a la investigación educativa, exponiendo, finalmente, cual es el marco teórico en el que se incluye.

1. CONCEPTO

Podríamos definir **IAO** como un grupo de programas informáticos, independientes o integrados en un paquete más heterogéneo de medios (libros, audiovisuales, programas pedagógicos, etc.) de diseño muy definido, el cual tiene la misión de analizar el comportamiento de los sujetos ante dicho programa. De esta definición se desprende que una **IAO** es una doble herramienta:

- herramienta para la **intervención**: por una parte, se presenta al usuario bajo un enfoque instructivo (por ejemplo, cualquier programa que enseñe a sumar, restar, leer, etc.).
- herramienta para la **investigación**: la otra faceta consiste en que mientras el usuario ejecuta el programa, paralelamente al aprendizaje, dicho programa realiza un *diario de datos*, que consiste en la recogida de todos los acontecimientos que suceden en la interacción del sujeto con el ordenador, en el que se refleja la conducta del usuario (teclas pulsadas, respuestas correctas e incorrectas, tiempo de latencia de la respuesta, etc.). Este diario ha de obedecer al diseño de la investigación y a la recogida de las variables previamente definidas.

Veamos un ejemplo de **IAO** aplicada a un programa de educación. Se trata de un programa diseñado desde el *Área de Métodos de Investigación* y aplicado gracias a un convenio que se estableció con el **RACC** (*Real Automóvil Club de Cataluña*), dirigido a los conductores responsables pero que desconocen algunas medidas preventivas. Con este fin se envió un pequeño juego informático mediante el cual, ante determinadas imágenes y preguntas, adquirieran conciencia de la peligrosidad de cierto tipo de comportamientos en la conducción (velocidad inadecuada, adelantamientos indebidos,...). Definida de este modo nos hemos encontrado ante una **EAO** más, es decir, un paquete informático que realiza unas tareas de instrucción o, en este caso, ayuda o refuerza la acción de un programa educativo.

Sin embargo, y este es el rasgo característico de las **IAO**, a este programa informático se le añadió un módulo captador encargado de recoger la reacción de los conductores para su posterior análisis tanto cuantitativa como cualitativamente. Este análisis estadístico nos indicó qué factores de riesgo se podían detectar en la respuesta de los conductores tomados como muestra (téngase en cuenta que dicha muestra era bastante amplia y los datos se habían obtenido rápidamente sin necesidad de repartir cuestionarios, etc., y de forma bastante objetiva e imparcial, puesto que el observado recibía ante él un juego, más que un programa acusador, por lo que no era consciente de que sus datos quedaban recogidos para su posterior estudio).

Esta investigación refleja adecuadamente nuestra concepción de una **IAO** estándar; no obstante, presentamos dos **IAO** aplicadas a campos distintos realizadas en nuestro laboratorio: la primera, se aplicó en métodos de lectoescritura para niños con deficiencias (realizándose en colaboración con las entidades *ASPACE*), mientras que la segunda se aplicó a la enseñanza de las ciencias naturales en el bachillerato (donde participaron varios institutos de secundaria). La convergencia entre ambas radica en que el uso de la informática nos permite no sólo recoger variables apreciables (respuestas correctas e incorrectas, etc.) sino también controlar el tiempo en responder a las preguntas y ver si las diferentes conductas de respuesta nos indican individual o globalmente algo.

* **IAO «leemos juntos»:** El programa informático *leemos juntos*, se presentó ante individuos con deficiencias mentales y sirvió para que, a la vez que jugaban, aprendieran a leer y escribir mediante un método de lectura global. Del diseño instructivo se ocupó un didacta, mientras que los investigadores del área de métodos especificaron las variables a estudiar y como se analizarían los datos recogidos. Para cada sujeto se obtuvo un fichero (que podía ser exportado a SPSS o cualquier otro sistema de tratamiento estadístico) en el cual constaba la fecha de realización de la experiencia instructiva, las respuestas acertadas y su especificación, las respuestas falladas y su especificación, y los tiempos de respuesta empleados en cada una de ellas. En estos ficheros históricos se observaba la evolución de cada uno de los alumnos, aportando una gran fuente de feed-back entre el nivel de asimilación del alumno y el conocimiento que sobre este tiene el profesor de cara a las siguientes estrategias de instrucción. A nivel global, también se disponía de los mismos datos permitiendo la realización de comparaciones de grupos, establecer diferencias entre grupo control y experimental, etc.

* **TESCOM Programa de aprendizaje de las ciencias naturales a nivel de secundaria:** Esta **IAO** consistió en un programa informático sencillo en el que el usuario debía clasificar, según una definición previa, distintas especies animales. Esta investigación puso de manifiesto que los individuos que utilizaban este programa tendían a tener mejores notas, o mejor asimilación y comprensión de los conceptos biológicos que se exponían en él, que los alumnos que sólo recibían clase magistral. Para la discusión sobre la influencia de variables de tipo motivacional, apetitivo, etc, sobre estas diferencias, y también si desea conocer con mayor detalle los resultados obtenidos en dicha investigación, remitimos al lector a la bibliografía (BOU, 1991).

Por otra parte, a nivel de grupo y utilizando los *tiempos de comprensión* y *tiempos de respuesta*, se pudieron constatar varias teorías o enunciados cognitivos, que requerían de complejos diseños experimentales para su realización. En base a los *tiempos de comprensión*, por ejemplo, se pudo detectar tendencias en el grupo que hacían que el individuo tardase más en comprender ciertos conceptos que otros (se consideró *tiempo de comprensión* al espacio de tiempo transcurrido desde que aparecía la definición de la clase biológica a estudiar, hasta que el individuo pulsaba la tecla manifestando que ya lo había entendido y respondía a la pregunta). Estas correlaciones se hicieron de la forma *fuerte*, es decir, con el coeficiente de correlación *W* de Kendal, el cual demostró que las ordenaciones para ciertas familias de conceptos que hacían los individuos eran las mismas, o sea, que había uniformidad de criterios en ordenar de más difícil a más fácil ciertos conceptos biológicos. La ordenación, además, se hizo inconscientemente, puesto que el alumno sólo sabía que tenía que hacer el trabajo y responder

cuando creyera que ya lo sabía. El alumno no afirmaba a nivel consciente «esto es más fácil que esto otro», sino que para deducir dicha información se utilizaba el tiempo empleado en entender el concepto. Una forma *débil* de hacerlo, hubiera sido calculando las medias de los tiempos de comprensión y las desviaciones, y ordenarlos respecto a las medias, existiendo un tratamiento similar para los tiempos de respuesta.

Otra conclusión a la que se llegó es que no había concordancia entre los *tiempos de comprensión* y los *tiempos de respuesta*, apoyando ciertas teorías cognitivas que establecen que los mecanismos que se activan durante la comprensión o durante la resolución de problemas no son siempre los mismos o coincidentes.

Puesto que el programa informático constaba de veintiocho preguntas, para cada alumno se generaba un fichero de datos en el que se recogían, entre otros, el tiempo de comprensión del enunciado, el tiempo de respuesta al aparecer el animal a clasificar, la respuesta correcta si lo había sido y, en caso de error, la respuesta elegida. Se puede apreciar como el volumen de información recogido para cada individuo era inmenso, lo cual hubiese supuesto meses de trabajo de haberse realizado siguiendo la metodología tradicional. Sin embargo, con este método los datos quedaban automáticamente grabados en formato listo para exportar a SPSS para su posterior análisis estadístico.

2. VENTAJAS DE LAS IAO

En las investigaciones expuestas hemos podido observar como las **IAO** ofrecen, a nuestro entender, una serie de ventajas las cuales resaltamos a continuación:

- 1) *Recogida directa de datos*. Las **IAO** simplifican todo este proceso. Formalmente, esta tarea es una pérdida de tiempo puesto que el usuario plasma sus datos en un papel y es necesario que posteriormente un operador de informática los traspase al ordenador. ¿Por qué no hacer que sea el usuario quien directamente los introduzca en el ordenador? Este es uno de los objetivos de las **IAO**, así, a la vez que garantiza la completa fiabilidad de todo el proceso, lo hace más económico.
- 2) *Control de errores*. La recogida de datos de las **IAO** elimina fuentes de error, dado que no pueden cometerse errores involuntarios de transcripción de datos, ni errores intencionados debido a la voluntad del usuario que es objeto de la investigación.
- 3) *Estudio de campo*. Las **IAO** siempre son investigaciones *vivas*, elaboradas en contextos externos al laboratorio, situaciones reales en las que el sujeto actúa de forma natural al no sentirse evaluado.
- 4) *Multienfoque*. En las **IAO** intervienen varias modalidades de investigación, permitiendo que los distintos equipos (investigación, orientación, didáctica,...) puedan colaborar en proyectos donde cada uno de ellos tiene tareas complementarias y definidas.

3. EQUIPO DE INVESTIGACIÓN

Detrás de una **IAO** siempre debe existir un *equipo de investigación*, es decir, un grupo mínimo de personas para que puedan realizarse investigaciones que contengan una **IAO**. En algunos casos una misma persona puede asumir varias funciones, pero en general para que una **IAO** sea viable es necesario contar con:

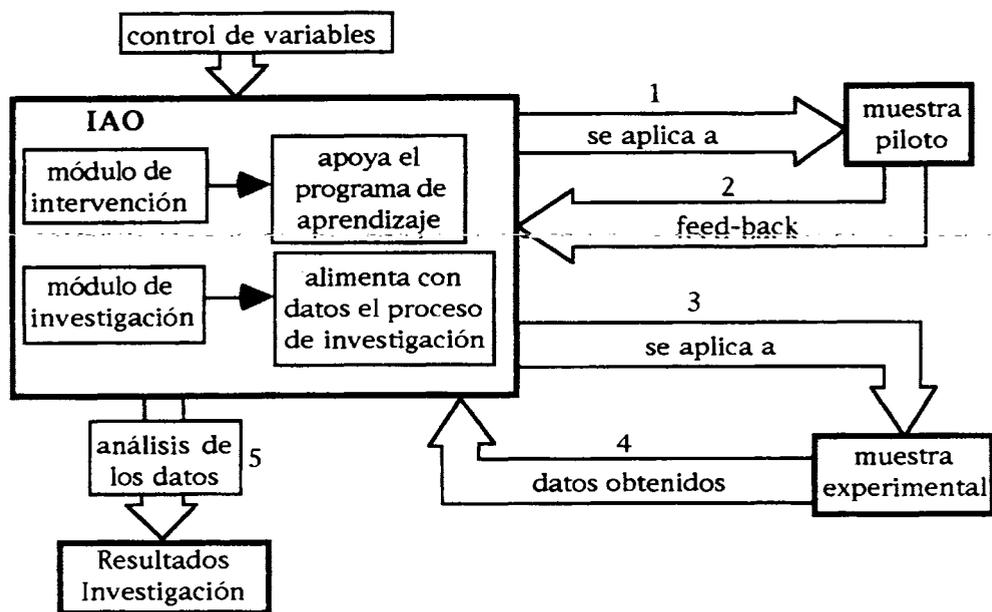
- a) un *especialista en didáctica, orientación* o en general conocedor del destinatario, el cual establecerá las pautas del programa, los contenidos educativos y los niveles más adecuados, su presentación y sus formas de intervención.
- b) un *investigador familiarizado con la investigación educativa*, encargado de establecer las

variables dependientes y independientes a estudiar, el tipo de control necesario para estas variables, la existencia o no de variables extrañas que puedan interferir en los resultados obtenidos y, en definitiva, todo el diseño cuasiexperimental.

- c) un *equipo técnico* adicional, en general suelen ser informáticos que colaboran en la realización de la IAO, grafistas especializados en el dibujo, etc.
- d) el *controlador de contenidos* o persona especialista en la materia que se investiga, la cual incide en los aspectos a profundizar y qué interesa recoger.

4. MARCO TEÓRICO

Las IAO aportan, por tanto, un nuevo marco de referencia de investigación que es el que resumimos en el siguiente cuadro, donde podemos apreciar como en una fase inicial se realiza la captación de datos en muestras pequeñas, llamadas *muestras piloto*, de las que se extraen las primeras conclusiones y directrices. Después, la IAO es aplicada a grandes muestras, *experimentales*, consiguiendo la captación definitiva de datos que nos permitan, mediante el consecuente análisis estadístico, conocer los resultados obtenidos en la investigación. Téngase en cuenta que de este modo es posible el refinamiento entre la primera y segunda fase, de forma que los posibles errores existentes en la investigación se pulen al analizar los resultados de las muestras piloto sin perturbar los resultados de la misma.



Este esquema abre las puertas a la investigación cualitativa y cuantitativa casando ambos paradigmas, dado que la primera fase de la investigación tiene un carácter más cualitativo mientras que la segunda es más cuantitativa. Esta primera fase debe ser sometida a las pruebas estadísticas habituales (fiabilidad de cuestionarios, resultados, Alfa de Cronbach y todas las que ya conocemos) permitiendo la realización de un análisis cualitativo más minucioso (indicándonos posibles deficiencias, dobles sentidos o preguntas ambiguas). No hace falta decir que este esquema sería tremendamente costoso si utilizamos un diseño de investigación clásico.

Laboratorio de Informática Aplicada a la Educación, Departamento de Pedagogía Aplicada, Facultad de Educación, Universidad Autónoma de Barcelona. 08193 Bellaterra.

BIBLIOGRAFÍA

BOU, G. (1991): *Aprenentatge comprensiu i processos d'informació: aplicacions educatives*. Universidad Autónoma de Barcelona, Tesis Doctoral.

NORMAN, D. A. (1982): *El aprendizaje y la memoria*. Alianza, Madrid.

[Nota: debido a problemas de extensión de la comunicación remitimos exclusivamente las obras citadas. Para cualquier información pueden contactar con el Laboratorio de Informática Aplicada a la Educación.]

RELACIONES INTERPERSONALES EN CENTROS EDUCATIVOS

Juana M^a Cruz Martínez y Mercedes Martín Berrido
Universidad de Sevilla

1. INTRODUCCIÓN AL ÁREA PROBLEMÁTICA

La **Teoría Tridimensional de la Personalidad** desarrollada por William Schutz (1966, 1967, 1978, 1979) se basa en que las personas en su interrelación con los demás se mueven a partir de tres necesidades psicológicas básicas: la **inclusión**, el **control** y el **afecto**. La inclusión es la primera necesidad que una persona manifiesta hacia los demás, se refiere a la necesidad de ser tenido en cuenta, de ser considerado por el otro. El control es la segunda necesidad básica y se refiere al poder que ejercemos sobre el comportamiento de los demás o al poder que dejamos en los demás con respecto a nuestro comportamiento propio. El afecto (o afección), tercera necesidad básica y nivel más cercano de relación interpersonal, se refiere al sentimiento de cariño (amor-odio), de sentirse querido y querer a otros. Estas dimensiones siempre se mueven en dos polos opuestos.

El equilibrio entre las tres dimensiones, entre los polos y colectivos, permitirá relaciones interpersonales sanas y satisfactorias. Los comportamientos de los individuos varían según hacia el polo en que se sitúe en cada dimensión y así, en cada una encontramos un comportamiento deficitario, excesivo, patológico o ideal; este último se conseguirá en el caso que todas las necesidades interpersonales básicas estén satisfechas.

Las aplicaciones de esta Teoría a los grupos escolares (aulas, centros y comunidades educativas) ayuda a comprender las relaciones grupales, por lo que la orientación de grupos puede tener en la Teoría de Schutz importantes aportaciones para su adecuada realización. A tal efecto, Pérez Juste (1990) advierte que las **relaciones interpersonales** ocupan un lugar privilegiado como elemento favorecedor de la armonía interior, de la adaptación o de los desequilibrios y conflictos, representando un evidente papel condicionante sin cuya alteración muchos problemas educativos (por tanto personales) no podrán ser resueltos.

2. METODOLOGÍA DE LA INVESTIGACIÓN

La investigación realizada aborda el diagnóstico de necesidades (**needs assessment**) ligadas a las relaciones interpersonales que se manifiestan en el contexto educativo. Para nuestro trabajo, entendemos por necesidad aquella discrepancia entre la situación manifiesta y la deseada. Estas discrepancias se han determinado en una doble vertiente: desde una **perspectiva democrática** (Stufflebeam, 1984) o cambio deseado por las audiencias implicadas y desde una perspectiva de **relación**, que resulta de comparar distintas situaciones o distintos grupos (Tejedor, 1990).

2.1. Los objetivos del estudio

En la línea de las diferentes investigaciones que han venido verificando esta teoría sobre relaciones interpersonales en orientación (St. Clair, 1961; Checkon y otros, 1972; Weiner, 1974; Schutz, 1977; Bueno y Castanedo, 1990; Castanedo y Bueno, 1990; entre otras), hemos desarrollado los siguientes objetivos:

- a) Describir las relaciones interpersonales entre profesores y alumnos, en el marco de la teoría expuesta.
- b) Analizar y comparar las necesidades que mantienen dichos alumnos y profesores.

2.2. Población y muestreo

La muestra utilizada ha sido de 67 profesores y 942 alumnos de 6º, 7º y 8º de E.G.B. distribuidos en centros que se adaptan a la características que se especifican en la tabla 1.

TABLA 1

CENTRO		TIPO		ZONA		INTEGRACIÓN	
RURAL	URBANO	PÚBLICO	PRIVADO	CAEP	NO CAEP	SI	NO
433	578	863	148	171	839	302	708

2.3. Procedimientos de recogida de datos

Hemos utilizado la versión facilitada por Castanedo (1990) de las siguientes escalas del FIRO (Fundamental Interpersonal Relations Orientation): VAL-ED (Educational Values), sobre valores educativos; FIRO-B (Fundamental Interpersonal Relations Orientation Behavior), sobre conducta interpersonal para profesores; y FIRO-BC (Fundamental Interpersonal Relations Orientation Behavior Childrens), para alumnos.

3. RESULTADOS Y CONCLUSIONES

3.1. Necesidades de los diferentes colectivos

Se aplicó la prueba T-Test Pairs que, a un nivel de significación de $\alpha=0,05$, nos permitió conocer cuáles eran las diferencias entre las conductas **expresadas** y **deseadas**.

TABLA 2

VARIABLES	T	α	Ho
INCLUSIÓN EXPRESADA-INCLUSIÓN DESEADA	9,74	0,000	NO ACEPTA
CONTROL EXPRESADO-CONTROL DESEADO	14,57	0,000	NO ACEPTA
AFECTO EXPRESADO-AFECTO DESEADO	2,38	0,017	NO ACEPTA

Los resultados obtenidos en el FIRO-BC que se presentan en la tabla 2, indican que la conducta de los alumnos está **muy desequilibrada**, ya que lo deseado discrepa significativamente de las conductas expresadas. Esto significa que los alumnos tienen una actuación en el entorno escolar que no se corresponde con sus deseos reales de actuación en ninguna de las tres necesidades básicas a las que nos referimos.

TABLA 3

VARIABLES	T	a	Ho
INCLUSIÓN EXPRESADA-INCLUSIÓN DESEADA	5,65	0,000	NO ACEPTA
CONTROL EXPRESADO-CONTROL DESEADO	0,89	0,376	SE ACEPTA
AFECTO EXPRESADO-AFECTO DESEADO	2,49	0,016	NO ACEPTA

En los resultados obtenidos para el colectivo de profesores que se muestran en la tabla 3, el desequilibrio sólo se da en las variables **inclusión** y **afecto**, no encontrándose diferencias significativas con respecto al control. Estos resultados desvelan que la inclusión real de los profesores no es la que ellos desean, al igual que ocurre con las relaciones afectivas que mantienen dentro de la comunidad escolar. Sin embargo, el control que están ejerciendo en la escuela se corresponde con el deseado.

3.2. Comparación intergrupal

Para comparar la opinión entre profesores y alumnos sobre valores educativos (VAL-ED), se aplicó asimismo la prueba T de STUDENT, con un $\alpha=0.05$, cuyos resultados pueden observarse en la tabla 4. Como podemos ver, existen diferencias significativas:

- Mientras que los alumnos opinan que no hay necesidad de colaboración entre profesores y los directivos para el diseño de actividades, los profesores opinan lo contrario.
- Según los alumnos, el equipo directivo debe controlar al profesorado, mientras que éste piensa que no.
- Los alumnos piensan que la comunidad debe controlar las actuaciones de los profesores, negándose éstos a tal control.
- Los alumnos piensan que los padres no tienen que tomar decisiones en la escuela, mientras que los profesores opinan que la intervención de los padres podría ser buena.
- El afecto entre profesores y alumnos es un factor que tanto unos como otros ven importante, sin embargo, el profesorado da puntuaciones extremas en esta variable lo que nos indica que lo consideran de vital importancia.
- El afecto entre Equipo directivo y profesores debe estar ligado a gustos personales según los profesores; sin embargo, los alumnos piensan que el profesorado debe relacionarse con el equipo directivo aunque no sean afines.

Como puede apreciarse en la tabla 4, las opiniones sobre el resto de las variables no arrojan diferencias significativas entre alumnos y profesores, por lo que podemos considerar que la instituciones se encuentran equilibradas respecto de las mismas.

TABLA 4

VARIABLES	T	a	Ho
INCLUSIÓN EQUIPO DIRECTIVO-PROFESORES	0,85	0,000	NO ACEPTA
INCLUSIÓN PROFESORES-PADRES	1,72	0,089	SE ACEPTA
CONTROL PROFESORES-ALUMNOS	0,33	0,743	SE ACEPTA
CONTROL EQUIPO DIRECTIVO-PROFESORES	6,03	0,000	NO ACEPTA
CONTROL PROFESORES-PADRES	5,55	0,000	NO ACEPTA
CONTROL EQUIPO DIRECTIVO-PADRES	9,59	0,000	NO ACEPTA
AFECTO PROFESORES-ALUMNOS	3,49	0,000	NO ACEPTA
AFECTO EQUIPO DIRECTIVO-PROFESORES	1,72	0,000	NO ACEPTA
AFECTO PROFESORES-PADRES	0,91	0,366	SE ACEPTA
AFECTO EQUIPO DIRECTIVO-PADRES	1,74	0,086	SE ACEPTA

ÚLTIMAS APORTACIONES EN LA EVALUACIÓN DEL PROGRAMA DE ORIENTACIÓN METACOGNITIVA DE LA COMPRESIÓN LECTORA (DRA. E. REPETTO Y COL.)

Elvira Repetto

Directora

Narciso Barrero, Roberto Castro, Juan Antonio Gil, Concha Martín

Colaboradores

INTRODUCCIÓN

Entre las aportaciones de la evaluación de programas de intervención aplicados que se aplican en nuestro país, merece destacarse por la novedad de su temática y la relevancia de sus resultados la evaluación del programa de orientación metacognitiva de la comprensión lectora (OMECOL), de la profesora Elvira Repetto (1989, 1990), la cual desde 1987 inició la línea de investigación sobre la orientación metacognitiva dentro de un modelo de aprendizaje mediado.

En los últimos años se han venido aplicando programas cognitivos dependientes de las tareas como el Programa de Enriquecimiento Instrumental (Feuerstein, 1980) y el Plan Harvard (Bolt, Beranek y Newman, 1984). Sin embargo el citado programa OMECOL constituye una novedad por su base metacognitiva, por su dependencia de las tareas escolares y por su presentación en formato multimedia.

En la presente mesa redonda vamos a sintetizar los elementos básicos del programa y los resultados de uno de los análisis estadísticos para caracterizar las ganancias relativas obtenidas con la aplicación del programa.

La mesa redonda se estructura en cinco apartados:

1. Resumen de la fundamentación teórica en la que se basa la línea de investigación de orientación metacognitiva dentro de una concepción de aprendizaje mediado.
2. Diferencia entre los programas cognitivos aplicados en el ámbito hispano y el programa OMECOL.
3. Finalidades del programa y enfoque estratégico.
4. Análisis de algunos resultados estadísticos.
5. Conclusiones de la investigación.

1. RESUMEN DE LA FUNDAMENTACIÓN TEÓRICA

La mejora de la lectura mediante programas metacognitivos estándar es un paradigma con entidad propia, habiéndose creado en numerosas universidades, sobre todo en el ámbito anglosajón, departa-

mentos específicos, integrados en la I.R.A. (International Reading Association), (Pagés y Echevarría, 1989).

El Programa OMECOL en su creación, diseño y aplicación forma parte de la línea de investigación dirigida por la Dra. Elvira Repetto en su cátedra, denominada «La Orientación Metacognitiva en un modelo de aprendizaje mediado» (Repetto, 1990). La línea de investigación se inscribe en un paradigma de cognición del alumno y de mediación de la enseñanza, con un modelo tetraedral limitado por las características del aprendiz, demandas de la tarea, naturaleza de los materiales y procedimientos instrumentales (Baker y Brown, 1984). Este paradigma mediacional centrado en el alumno (Shulman, 1989) constituye un puente entre las concepciones proceso-producto y la ecología del aula, constituyendo una fusión de lo cognitivo con los aspectos motivacionales.

2. DIFERENCIA ENTRE LOS PROGRAMAS COGNITIVOS Y EL PROGRAMA OMECOL

Como el rendimiento se incrementa preferentemente dentro del área en que se ha trabajado y en función de las estrategias específicas implementadas (Stevens, 1988), los programas que son independientes de contenido tienen dificultades a la hora de la transferencia al currículo. Así otros programas como el P.E.I. (Fensterstein, 1990) construyen en los sujetos una matriz mental en léxico, conceptos, operaciones y relaciones; esta prótesis cognitiva actúa ortopédicamente, incrementando factor «g», pero la correlación entre «g» y rendimiento escolar ($=0.50$) aunque significativa no es muy elevada con lo que los alumnos mantienen déficits en cuanto a las materias de estudio y en cuanto a sus dificultades en dominios básicos instrumentales. Debido a esto existen programas cognitivos que para superar este problema de transferencia, se denominan como metacognitivos, cuando en realidad no son más que programas cognitivo-procesuales independientes de la tarea y sin estructura de auto-preguntas estratégicas. Los programas metacognitivos se caracterizan así porque se refieren a la auto-conciencia que el sujeto tiene acerca de su propio conocimiento.

3. FINALIDAD DEL PROGRAMA OMECOL

Este programa de orientación metacognitiva tiene como objetivo incrementar en los alumnos diversas variables pedagógicas y psicológicas, mediante actividades de autoobservación de la comprensión lectora que incrementan la auto-conciencia de los procesos cognitivos empleados.

El enfoque del Programa OMECOL es básicamente estratégico y se estructura fundamentalmente mediante auto-preguntas acerca de: los objetivos lectores y organizadores previos, adecuación de la velocidad, activación de predicciones, atención a lo relevante, atención a los tópicos, valoración de los conceptos, representación imaginativa, ayuda por fuentes diversas, re-lectura, lectura cooperativa y resúmenes.

4. ANÁLISIS DE ALGUNOS RESULTADOS ESTADÍSTICOS

En primer lugar vamos a relacionar las variables estadísticas utilizadas en el presente trabajo, así como los tratamientos realizados, circunscribiéndonos en este último punto al análisis de segmentación de la muestra en grupos en relación a diversas variables significativas para nuestro estudio y tomando como criterio variables de clasificación y otras relevantes, que posteriormente detallaremos. No es objeto de esta comunicación detallar tópicos relativos al diseño del experimento, muestreo o a instrumentos utilizados (OMECOL, CML), aún cuando estos últimos han sido reseñados anteriormente.

Las variables utilizadas se pueden agrupar en cuatro bloques:

1. Variables cognitivas:
 - a) Inteligencia general.
 - b) Aptitud verbal.
 - c) Razonamiento abstracto.
2. Variables pedagógicas:
 - a) Aprendizaje en CC. Sociales.
 - b) Vocabulario.
 - c) Comprensión lectora.
 - d) Conocimiento metacognitivo.
 - e) Conciencia de actividades lectoras.
3. Variables sociales y ocupacionales:
 - a) Status sociocultural.
 - b) Ocupación de los padres.
4. Variables de clasificación:
 - a) Localidad.
 - b) Comunidad Autónoma.
 - c) Zona.
 - d) Curso.
 - e) Sexo.
 - f) Edad.
 - g) Centro

El objetivo perseguido con este estudio de segmentación muestral, es como su nombre indica, la formación de grupos homogéneos que me permitieran caracterizar las ganancias relativas (diferencia entre puntuaciones pre-pos test dividida entre puntuación pretest) en variables como conocimiento metacognitivo, conciencia de actividades lectoras o comprensión lectora, en las que intervenía el programa OMECOL, en función de las restantes variables.

Para conseguir el objetivo definido se ha utilizado el programa CHAID (Chi-squared Automatic Interaction Detection). Esta técnica, desarrollada por Cellard et al. (1967), Bourouche y Tennenhaus (1972), Kass (1980) y Magidson (1989), quien la ha adaptado para el SPSS, tiene como principal característica el que la variable dependiente es categórica, es un sistema de clasificación basado en la estructura de árbol que deriva su justificación de las tablas de contingencia estadísticas y que ha demostrado su aplicación fundamental en la segmentación de mercados.

Dejamos para la mesa redonda el entrar en detalles sobre esta técnica de segmentación y vamos a continuación a presentar de forma esquemática alguno de los resultados obtenidos y las conclusiones de los mismos.

5. CONCLUSIONES DE LA INVESTIGACIÓN

En el presente trabajo, se han realizado tres estudios de segmentación que se diferencian en la variable dependiente utilizada: ganancias relativas en comprensión lectora (ngacolec), ganancias relativas en la conciencia de actividades lectoras (ngacml) y finalmente, ganancias relativas en conocimiento metacognitivo (ngaome), en *anexo* presentamos los resultados de esta última variable, para identificación de los mismos en la *tabla 1* establecemos un paralelismo entre la denominación de la variable en el estudio con su significación.

TABLA 1
IDENTIFICACIÓN DE LAS VARIABLES

Dependent Variable	Niveles			Categoría		Frequency Counts
				#	Label	
ngaome (Conocimiento metacognitivo)	2			1	Cat. 01	850
				2	Cat. 02	829
					Total	1679
Predictor	Niveles	Combine Type	Sig	#	Categoría Sym:Label	Frequency Counts
com	3	Free	0.05	1	1:ANDALUCÍA	657
				2	2:CANTABRIA	269
				3	3:GALICIA	753
ni	3	Free	0.05	1	1:BÁSICO	461
				2	2:MEDIO	375
				3	3:SUPERIOR	843
gr	2	Free	0.05	1	1:EXPERIMENTAL	839
				2	2:CONTROL	840
zo	3	Free	0.05	1	1:URBANA	608
				2	2:SEMIURBANA	362
				3	3:RURAL	709
cu	7	Free	0.05	1	1:SEGUNDO	218
				2	2:TERCERO	242
				3	3:CUARTO	189
				4	4:QUINTO	187
				5	5:SEXTO	319
				6	6:SÉPTIMO	245
				7	7:OCTAVO	279
edad	9	Free	0.05	1	1:SIETE	197
				2	2:OCHO	249
				3	3:NUEVE	146
				4	4:DIEZ	199
				5	5:ONCE	242
				6	6:DOCE	264
				7	7:TRECE	259
				8	8:CATORCE	107
				9	9:QUINCE	16

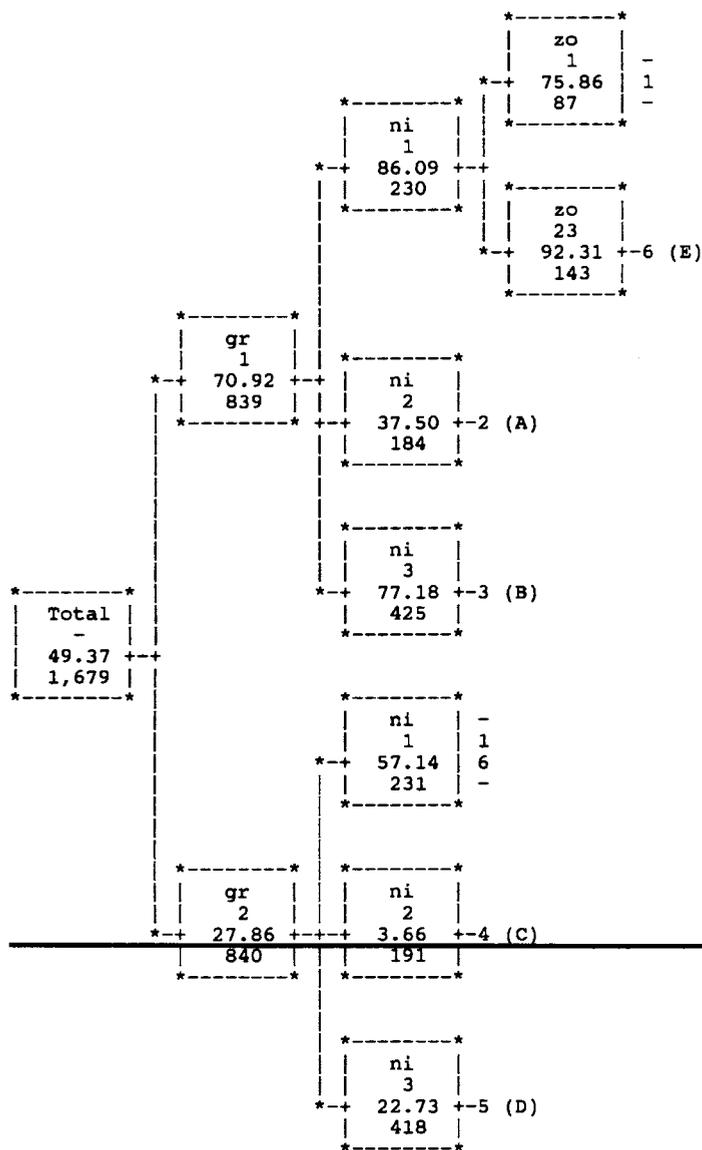
Predictor	Niveles	Combine Type	Sig	#	Categoría Sym:Label	Frequency Counts
sexo	2	Free	0.05	1	1:HOMBRE	872
				2	2:MUJER	807
nop (Ocupación de los padres)	2	Free	0.05	1	1:Cat. 01	1038
				2	2:Cat. 02	641
nsc (Status socio-cultural)	2	Free	0.05	1	1:Cat. 01	854
				2	2:Cat. 02	825
ngafacg (Inteligencia general)	2	Free	0.05	1	1:Cat. 01	835
				2	2:Cat. 02	844
ngaapver (Aptitud verbal)	2	Free	0.05	1	1:Cat. 01	844
				2	2:Cat. 02	835
ngaraabs (Razonamiento Abstracto)	2	Free	0.05	1	1:Cat. 01	832
				2	2:Cat. 02	847
ngaprsoc (Aprendizaje en CC.Sociales)	2	Free	0.05	1	1:Cat. 01	840
				2	2:Cat. 02	839
ngaprvc (Vocabulario)	2	Free	0.05	1	1:Cat. 01	841
				2	2:Cat. 02	838
ngacolec (Comprensión lectora)	2	Free	0.05	1	1:Cat. 01	845
				2	2:Cat. 02	834
ngacml (Conciencia de actividades lectoras)	2	Free	0.05	1	1:Cat. 01	851
				2	2:Cat. 02	828

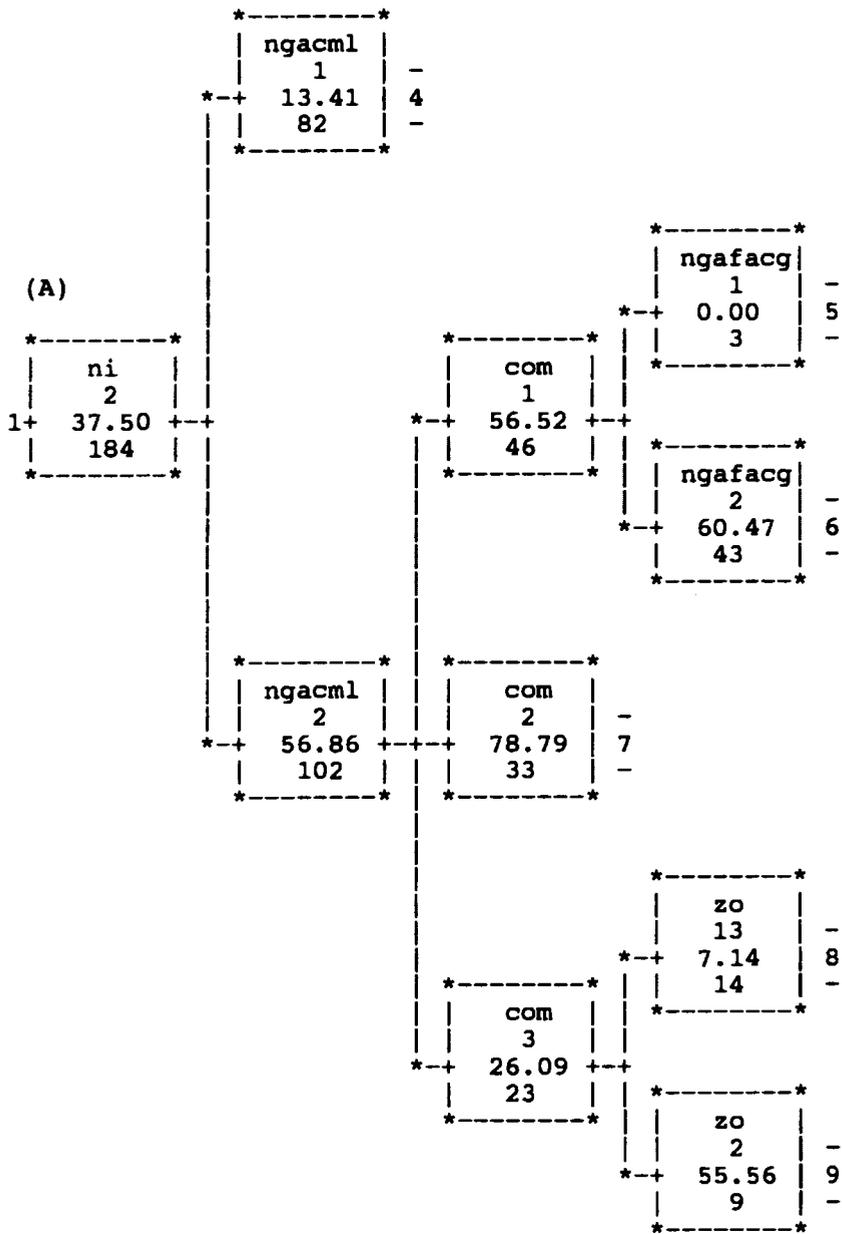
Como consecuencia de los resultados que hemos indicado las principales conclusiones que podemos obtener son las siguientes: a) Los alumnos que tienen una ganancia relativa en conocimiento metacognitivo (ngaome) superior a la mediana, proceden del grupo experimental y son alumnos del ciclo superior. Representan un 20% del total de los 1.679 alumnos de la muestra.

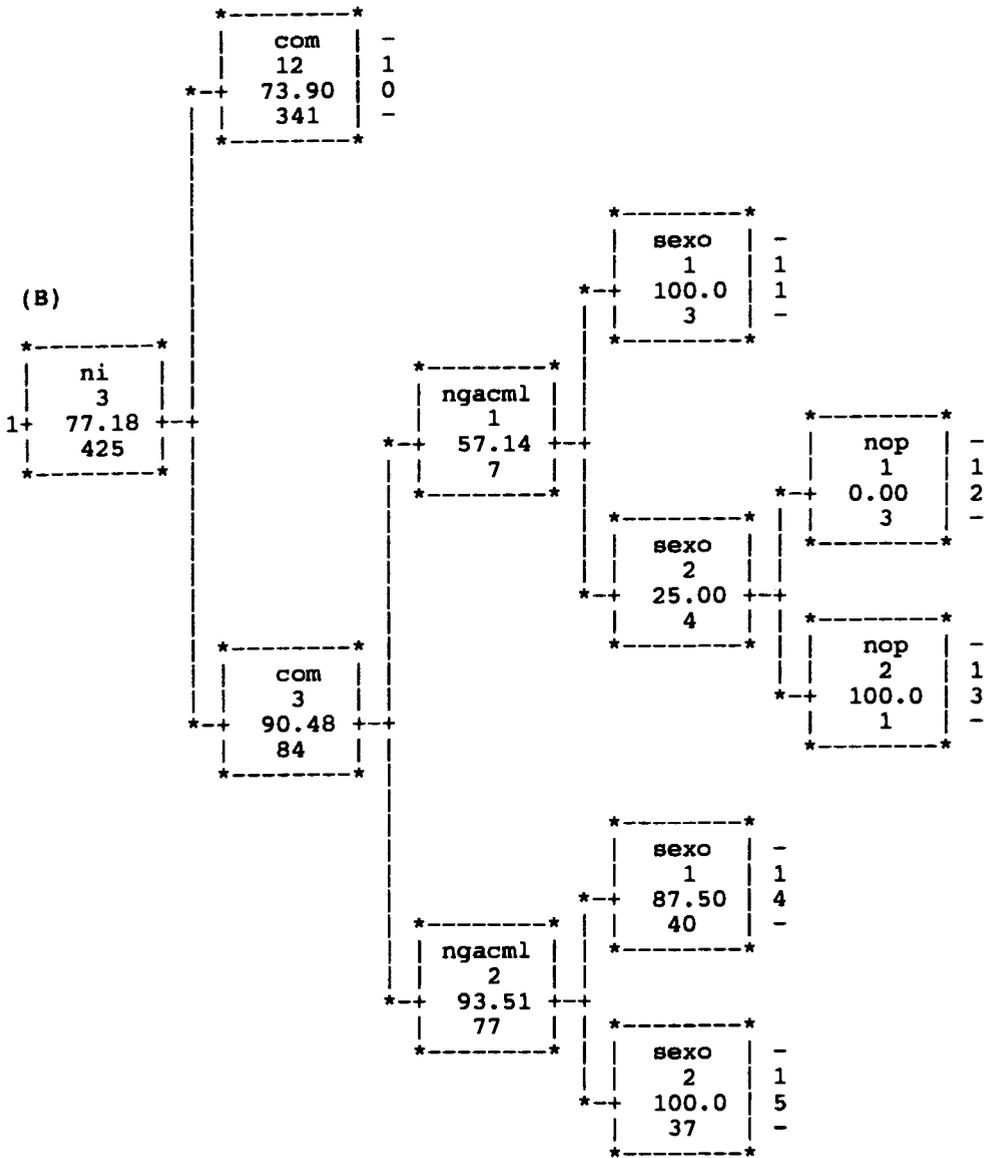
b) Del grupo anterior el 77% proceden de las Comunidades de Andalucía y Cantabria y el resto de Galicia. Este último grupo es de mujeres con una ganancia relativa alta (mayor que la mediana) en conciencia de actividades lectoras.

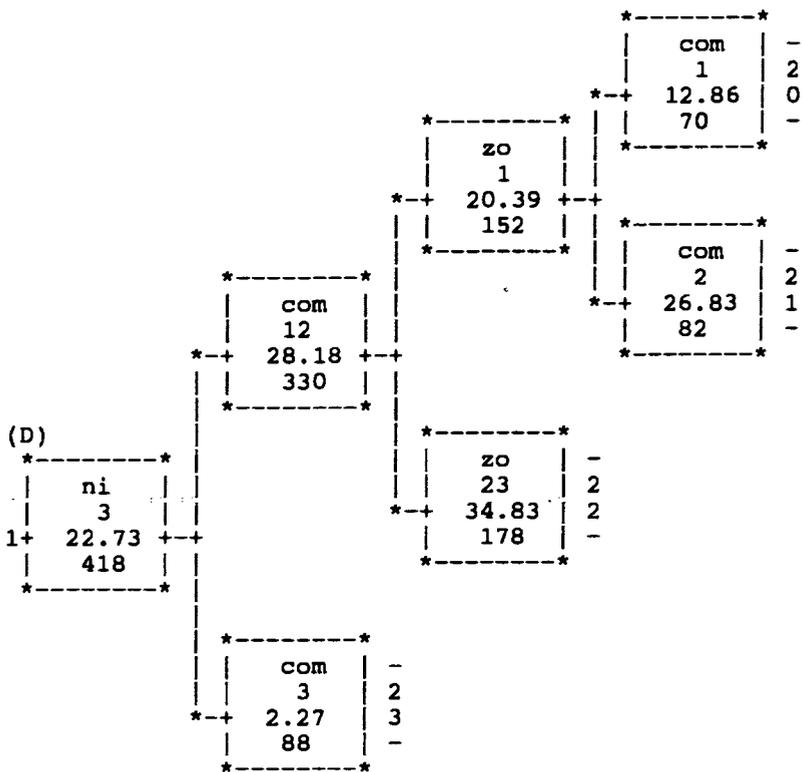
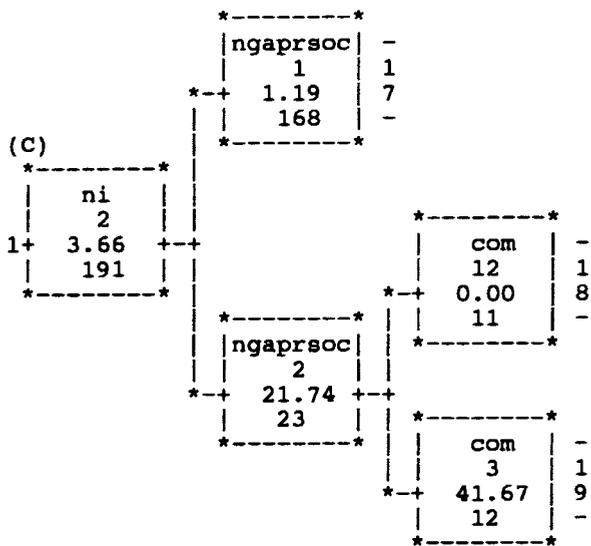
c) Existe un grupo pequeño de un 5% del total, alumnos del ciclo superior y que proceden del grupo control.

ANEXO









(E)	zo 23	com 12	-
	92.31	0.00	2
1+	143	1	-
		com 3	-
		92.96	3
		142	-

BIBLIOGRAFÍA

- BAKER, L. y BROWN, A. L. (1984): «Cognitive monitoring in reading». J. FLOOD (Ed.): Understanding Reading Comprehension. N. York: D.E. (IRA).
- BOLT, BERANEK y NEWMAN (1984): Project intelligence. Harvard: Univ. Press.
- ESCOBAR, M. (1992): «El análisis de segmentación, concepto y aplicaciones. Estudio/Working Paper 1992/31. Instituto Juan March de Estudios e Investigaciones. Madrid.
- FEUERSTEIN (1980): «Instrumental enrichment. An intervention program for cognitive modifiability». Baltimore: Univ. Park Press.
- PÁGES y ECHEVARRÍA (1989): «Investigación internacional sobre la lectura». Madrid. Asociación SM. Madrid.
- REPETTO, E. (1987): *Programa «OMECOL»*. Madrid: UNED (Cátedra de Orientación).
- REPETTO, E. (1989): «Intervenciones orientadoras y currículum educativo». Jornadas de Orientación Escolar y Profesional. Las Palmas: AEOEP.
- REPETTO, E. (1990): «El entrenamiento metacognitivo, la modificación cognitiva y su transferencia a la comprensión lectora, a la resolución de problemas y al aprendizaje». *Revista de Investigación Educativa*, 8, (16), 563-587.
- STEVENS, R. J. (1988): «Effects of strategy training on the identification of the main idea of expository passages». *Journal of Educational Psychology*, 80, (1), 21-6.
- TIERNEY, R. J. (1983): «Learning from text». Reading Educational Report, nº 37.

EVALUACIÓN PROCESUAL DEL PROGRAMA HELIOS. UN ESTUDIO CUALITATIVO DE CASO MÚLTIPLE¹

Samuel Fernández Fernández

Universidad de Oviedo

Ángel Lázaro Martínez, Pedro Muncio Fernández

Universidad Complutense

Isabel Torroba Arroyo

Ministerio de Educación y Ciencia

INTRODUCCIÓN

La finalidad de esta comunicación es dar a conocer el desarrollo evaluativo de un programa promovido por la Comunidad Europea y al que se adscriben diez proyectos españoles de características muy diferenciadas.

El programa, denominado HELIOS II (C.E., 1988), plantea una acción global sobre la minusvalía en el ámbito europeo, de forma que todos los proyectos participantes presentan acciones modelo, a difundir tanto en el marco local, como en el nacional y europeo. Estos proyectos se organizan en grupos o redes de rehabilitación y formación profesional (Red de Centros), de integración laboral (Red Económica), de integración social (Red Social) y de integración educativa (Red Educativa). (Norberto Fernández, 1990).

De acuerdo con las directrices de la CE, la calidad de la aportación de los proyectos requiere una supervisión eficaz que garantice que cada Acción Modelo Local (denominación comunitaria de los proyectos) desarrolla las iniciativas planteadas ante la CE. Por su parte, la dependencia de las tres redes (Centros, Económica y Social) del Instituto de Servicios Sociales (INSERSO), le permite a esta institución definir la utilidad que estas acciones o proyectos tienen en el marco de una política en favor de las personas discapacitadas.

OBJETIVOS DE LA EVALUACIÓN

Los principios generales del Programa HELIOS, relativos al desarrollo de un enfoque comunitario y basados en las experiencias de los países miembros sobre el fomento de la formación y rehabilitación profesional y la integración económica y social de los minusválidos, plantean la necesidad de desarrollar un sistema de evaluación que permita, tanto a los propios participantes, como a las instituciones de que dependen, conocer el proceso y los ajustes requeridos en función de los fines iniciales, así como desarrollar técnicas que determinen la validez y eficacia de las distintas acciones.

¹ El Proyecto de Evaluación del Programa HELIOS y el desarrollo de SEPIR, mencionados en este escrito, han sido financiados íntegramente por el Instituto de Servicios Sociales, INSERSO.

Partimos, además, del hecho de que los proyectos han sido elaborados sin una pauta común y se encuentran en pleno desarrollo en el momento en el que se diseña su evaluación. Los objetivos de evaluación para las tres redes, de Centros, Integración Económica e Integración Social, están, por tanto, condicionados por la situación de los proyectos y se concretan en los siguientes:

1. Determinar la «situación actual» de cada uno de los proyectos y redes según una serie de parámetros que permitan establecer un marco de referencia comparativa intra e interproyectos.

2. Ayudar a los responsables de los proyectos a desarrollar las actividades y a conseguir los fines planteados mediante el análisis y la reflexión conjunta sobre los datos iniciales, las acciones en curso y los logros parciales.

Como objetivo suplementario, la actividad evaluadora se plantea el servir de base para el posterior desarrollo de un conjunto de instrumentos (Sistema de Evaluación de Programas de Integración y Rehabilitación, SEPIR, Fernández (1993), Lázaro (1993), Municio (1993) y Torroba (1993)) cuyo fin es el análisis externo y la autovaloración de procesos de intervención social en el ámbito de las personas discapacitadas.

BASES METODOLÓGICAS DEL PROYECTO EVALUATIVO

El proyecto de evaluación se realiza en dos fases. En la primera se dimensionan los proyectos, se clarifican las actividades y se establecen los procesos a seguir; en la segunda se analiza el grado de consecución de las acciones, los impactos y las percepciones logrados en los diferentes ambientes en que se incardinan los proyectos.

En este documento abordamos, exclusivamente, el planteamiento procesual de la evaluación por considerar que éste constituye la fase más relevante, a nivel metodológico, de todo el procedimiento evaluador. (La fase de evaluación de productos o «final» y la ulterior o complementaria, relativa al desarrollo del sistema de evaluación de programas de integración y rehabilitación, SEPIR, pueden conocerse a través de los correspondientes informes de investigación).

La incorporación de un proceso evaluador a una acción o proyecto ya iniciado limita la elección y la organización del diseño, puesto que la unión ideal entre programa y evaluación se efectúa cuando ambos procesos se desarrollan, de principio a fin, en mútua interacción (un procedimiento unificado puede seguirse en Fernández y col., 1991). No obstante, consideramos que los procedimientos desarrollados por el Instituto de Evaluación de Educación Especial de Austin, EE.UU. (Borich y Jemelka, 1982), y en concreto el modelo de «Evaluación Respondente» de Stake, resultan perfectamente viables y adecuados para la incorporación de la tecnología evaluadora procesual a los diez proyectos sociales ya en marcha. La idoneidad del modelo viene dada por presentar un esquema recurrente basado en dos elementos necesarios en nuestra investigación, como son la reflexión, a nivel interno y externo, sobre los propósitos del proyecto y situación actual de realización, y el constituir un proceso permanente de toma de decisiones que corrige puntualmente las desviaciones, subsana los errores y supera los obstáculos, reconduciendo el proceso hacia la mejor dirección estimada.

El modelo, por tanto, resulta de interés por cuanto nos permite, como evaluadores externos, interaccionar con el equipo director del proyecto y estimular la autorreflexión y la acción «respondente» siguiendo una estructura funcional equiparable al conocido modelo del reloj (Stake, 1975).

El modelo evaluador puede también conceptualizarse como un estudio evaluativo de casos (Stenhouse, 1990), al ser analizados en profundidad, y valorados, una serie de proyectos adscritos a un programa con el propósito de proporcionar a los responsables institucionales la información que les ayude a juzgar el mérito y el valor de dichos proyectos. Los escasos pero comunes elementos que caracterizan a una red permiten considerar, igualmente, que el esquema evaluador puede responder al concepto de estudio de caso múltiple indicado por Stake y Easley (1978).

Así pues, el hecho de analizar una situación natural con el propósito de comprender e interpretar, de forma global, la realidad de lo que acontece en el entorno de los proyectos, nos sitúa ante un «estudio de casos de orientación cualitativa» (Garanto, 1993; Burgess, 1984). En este estudio, tanto el esquema de análisis, de carácter inductivo y descriptivo, como el papel del investigador, que cobra un especial protagonismo como provocador de autoanálisis, responden a un planteamiento de investigación evaluativa plenamente cualitativa (Bogdan y Biklen, 1982, pp. 27-30; Corrie y Zaklukiewicz, pp. 125).

TRABAJO DE CAMPO

Aunque la investigación analítica es vista a veces como imprecisa en su diseño y desarrollo, procede generalmente en función de estadios bien delimitados. Así, nuestra primera etapa ha consistido en la toma de contacto con los diez proyectos, dispersos por todo el estado español. La familiarización con los participantes, los recursos y las actividades nos ha permitido, también, diseñar un guión básico de recogida permanente de información en respuesta al esquema del modelo evaluativo adoptado.

La observación, durante varios días, de las acciones puestas en marcha por los proyectos y las entrevistas semiestructuradas, no directivas, realizadas por dos miembros del equipo evaluador (ver Bechofer et al, 1984 en relación con la utilización de varios entrevistadores), con los participantes en el proyecto, tanto personas discapacitadas, como equipos directivos y técnicos, nos ha permitido ir documentando la realidad de todas las actividades, planificadas en un principio de forma un tanto heterodoxa.

La organización de la información constituye la tercera y más laboriosa de las etapas. En un primer momento se han desarrollado informes individuales y descriptivos de cada proyecto siguiendo un esquema de comparación de datos (Glaser y Strauss (1967) y un indexado tal y como recomienda Stenhouse (pp. 647)). Con posterioridad, la información ha sido resumida en tres informes narrativos, uno para cada red, en los que se recogen abundantes gráficas y matrices explicativas (Garanto, 1993; Miles y Huberman, 1987) que sintetizan los objetivos, procedimientos, variaciones, etc. de los tres tipos de proyectos.

CONCLUSIÓN

Los aspectos más relevantes de la evaluación procesual han permitido, por último, plantear un informe escueto de recomendaciones que ha sido presentado o remitido a todos los organismos e instituciones nacionales y europeas. Estas últimas síntesis han servido, también, para la elaboración de un instrumento que permitiese afrontar la evaluación final en la que se han valorado con posterioridad los productos e impactos logrados por las redes de centros de formación, de integración laboral o económica y de integración social, según un esquema matricial que recoge los logros a nivel individual y de repercusión local y nacional (micro, mesa o macro, en la denominación de Guffens o de cambio, coordinación y gerencia, en términos de Borich y Nance, 1987).

De acuerdo con la normativa del Joint Committee on Standards for Educational Evaluation, expresada en sus «Standars for Evaluation of Educational Programs, Projects and Materials» consideramos que la evaluación desarrollada en los distintos proyectos y redes del programa HELIOS cumple las cuatro bases de toda investigación evaluativa, que son:

1. *Ser útil*, lo que supone estar dirigida a aquellas personas y grupos relacionados con la tarea que se está evaluando, ayudándoles a identificar lo bueno y lo malo de su proyecto y a plantear soluciones para su mejora.
2. *Ser factible*, empleando procedimientos asequibles y controles que incidan positivamente en la evolución del proyecto.
3. *Ser ética*, por estar basada en compromisos explícitos de cooperación, protección de los derechos

de las personas (minusválidas en nuestro caso), honradez de los resultados y equidad en los informes.

4. *Ser exacta*, describiendo con claridad el proyecto en su contexto y desvelando las características de los procedimientos de la evaluación, por las que se aprecia que está libre de influencias y que las conclusiones son válidas y fidedignas.

REFERENCIAS BIBLIOGRÁFICAS

- BECHOFER, F., ELLIOTT, B. y MCCRONE, D. (1984): Safety in numbers: on the use of multiple interviewers. *Sociology*, 18, 1, 97-100.
- BURGESS, R. G. (1984): *The Research Process in Educational Settings: Ten Case Studies*. London: The Falmer Press.
- BOGDAN, R. y BIKLEN, S. K. (1982): *Qualitative Research for Education: An Introduction in Theory and Practice*. Hemel Hempstead: Allwyn and Bacon.
- BORICH, G. D. y NANCE, D. D. (1987): Evaluating special education programs: shifting the professional mandate from process to outcome. *RASE* 8 (3), 7-16.
- BORICH, G. D. y JEMELKA, R. P. (1982): *Programs and systems: An evaluation perspective*. New York: Academic Press.
- C.E. (1988): *Documento de Trabajo nº 7 de Evaluación*. Equipo de expertos de la Comunidad Europea. Bruselas. CEE.
- CORRIE, M. y ZAKLUKIEWICZ, S. (1985): Qualitative Research and Case-Study Approaches. En HEGARTY, S. y EVANS, P. (Eds.): *Research and Evaluation Methods in Special Education*. Windsor, NFER-NELSON.
- FERNÁNDEZ, S. (1993): *Guía de Evaluación de Programas de Integración Laboral*. Madrid: INSERSO.
- FERNÁNDEZ, S. CORRAL, M., MARTÍNEZ, R. A. y SAN FABIÁN, J. L. (1991): *Informe final de Evaluación del Proyecto Helios-Oviedo*. Oviedo: Dirección Provincial del MEC.
- GARANTO ALOS, J. y DEL RINCÓN IGEA, D. (1993): El Estudio de Casos. En GARCÍA PASTOR, C. (Ed.): *La Investigación en Educación Especial*. Salamanca: Amaru.
- GLASER, B. y STRAUSS, A. L. (1967): *The Discovery of Grounded Theory*. London: Weidenfield and Nicholson.
- JOINT COMMITTEE ON STANDARDS FOR EDUCATIONAL EVALUATION (1981): *Standards for Evaluation of Educational Programs, Projects and Materials*. New York: McGraw-Hill.
- LÁZARO, A. (1993): *Guía de Evaluación de Programas de Orientación Profesional*. Madrid: INSERSO.
- MILES, M. B. y HUBERMAN, A. M. (1987): *Qualitative Data Analysis*. Beverly Hills: Sage.
- MUNICIO, P. (1993): *Guía de Evaluación de la Gestión de Programas*. Madrid: INSERSO.
- NORBERTO FERNÁNDEZ, J. (1990): Dossier sobre redes comunitarias del programa Helios. *Minusval*, 70, pp. 13-42.
- STAKE, R. E. (1975): Program Evaluation: particularly responsive evaluation. Citado en Stufflebeam, D. L. y SHINKFIELD, A. J. (1987): *Evaluación Sistemática. Guía teórica y práctica*. Madrid. Paidós-MEC, pp. 262.
- STAKE, R. E. y EASLEY, J. (1978): *Case Studies in Science Education*. University of Illinois, Urbana, Illinois: CIRCE.
- STENHOUSE, L. (1990): Case Study Methods En WALBERG, H. J. y HAERTEL, G. D. (Eds.): *The International Encyclopedia of Educational Evaluation*. (1ª. ed.) Oxford. Pergamon.
- TORROBA, I. (1993): *Guía de Evaluación de Programas de Integración Comunitaria*. Madrid: INSERSO.

ESTUDIO DE LAS ACTITUDES DEL PROFESORADO HACIA LA INTEGRACIÓN DE LOS «ESCOLARES DIFERENTES» Y SU RELACIÓN CON LAS INTENCIONES DE CONDUCTA A PARTIR DEL MODELO DE AJZEN Y FISHBEIN

Luis Antonio Feliciano García
Universidad de La Laguna

INTRODUCCIÓN

Del mismo modo que en otras áreas del comportamiento, la importancia de las actitudes para los intercambios que se producen dentro del contexto escolar ha sido subrayada en numerosas ocasiones, por cuanto se estima que pueden promover o impedir el que estos tengan un carácter constructivo (Horne, 1985). Concretamente, en lo que se refiere al ámbito de **la integración de escolares con n.e.e.**, existe un acuerdo unánime en señalar a las actitudes que sostienen los profesores como uno de los factores que ejercen mayor influencia sobre su evolución; en la medida que estimulan o limitan la disposición a aceptar la incorporación de sujetos con n.e.e. en el aula ordinaria y la creación de un clima favorable a los mismos dentro de ella.

Este punto de vista ha llevado a la realización de múltiples investigaciones con el objetivo de determinar el tipo de actitudes que respecto al tema mantienen los citados agentes; siendo utilizadas para ello distintas técnicas de recogida de datos: Desde Escalas Likert a Listas de Control de Adjetivos pasando por cuestionarios, entrevistas, observaciones, etc. Ahora bien, dichos estudios, aparte de que suelen presentar una falta de concreción en el **objeto, acción, tiempo y contexto** de las actitudes analizadas, (dejándolos en muchas ocasiones a la libre interpretación del encuestado), carecen de una perspectiva sistémica que se ocupe no sólo de especificar el carácter positivo o negativo de esas valoraciones, sino también de examinar las relaciones que sostienen con otras variables como, p.e., las estructuras normativas y el campo de las intenciones de conducta de los individuos.

En base a ello, consideramos necesario el desarrollo de investigaciones que, mediante el examen de las citadas variables, nos proporcionen tanto una idea de sus características como del tipo de conexiones que existen entre ellas; con el objeto de poder determinar aquellos aspectos sobre los que se ha de intervenir a fin de que el profesorado se muestre más dispuesto a admitir a los alumnos con n.e.e. en el aula ordinaria.

Con este propósito, la investigación que a continuación vamos a comentar se ha fundamentado en los planteamientos de Ajzen y Fishbein (1980), quienes señalan que:

a) El determinante inmediato de toda conducta que se encuentra bajo el control volitivo de los individuos es la intención que estos tienen de realizarla o no.

b) La intención de conducta (BI) es una función de dos componentes: La actitud hacia la conducta (AB) y la norma subjetiva (SN).

c) La expresión algebraica de la relación entre la BI y sus componentes toma la forma de una ecuación de regresión múltiple.

d) La importancia relativa de la AB y de la SN en la predicción de BI puede ser estimada a través de los coeficientes obtenidos en el correspondiente análisis de regresión múltiple.

Teniendo en cuenta tales planteamientos, la investigación que a continuación pasamos a describir trata de: 1) Analizar el modo en que el profesorado de EGB valora la integración de 12 Alumnos-Problema; 2) examinar el tipo de normas que tiene al respecto; 3) Determinar si dichas actitudes y normas se hallan relacionadas; 4) comprobar si esas AB y SN se encuentran relacionadas con las intenciones de conducta de los enseñantes.

MÉTODO

En este estudio participaron 297 profesores de EGB en activo (217 mujeres y 80 varones), de los que 172 impartían clase en CS y 125 en PRESC, CI y CM.

Con el fin de obtener la información necesaria para la investigación se elaboró un cuestionario que contenía los siguientes apartados: a) Relativo a las actitudes de los encuestados hacia el hecho de integrar a tiempo completo¹ a doce Alumnos-Problema² (en el que la integración de cada alumno era valorada mediante un Diferencial Semántico formado por tres escalas de adjetivos bipolares); b) relativo a las normas subjetivas del profesorado sobre la integración de cada Alumno-Problema (en el que la probabilidad que un sujeto atribuye al hecho de que sus otros significativos piensen que debe integrar a un determinado alumno era examinada mediante una escala bipolar de siete puntos); c) Relativo a las intenciones de los encuestados respecto a integrar en sus clases a tiempo completo a cada Alumno-Problema (en el que la probabilidad que un sujeto atribuye al hecho de que quiera integrar a un determinado alumno era examinada mediante una escala bipolar de siete puntos).

El proceso de análisis de los datos comprendió: a) El examen de los estadísticos de cada variable; b) el examen de las relaciones entre las variables del modelo estudiado, mediante el coeficiente de correlación de Pearson y el análisis de regresión múltiple (método stepwise).

RESULTADOS

Por lo que respecta a los estadísticos de las AB analizadas, se comprobó que los promedios de las valoraciones realizadas por los profesores de la muestra acerca de la integración de cada Alumno-Problema diferían en función del tipo de deficiencia que estos presentaban; de tal modo que mientras los enseñantes se muestran bastante favorables a la integración de los alumnos más «acomodables» a la clase ordinaria (caso del Asmático Crónico, Diabético o Epiléptico), mantienen una postura desfavorable hacia la de los escolares menos adaptables (caso del Ciego, Sordo o Retrasado Mental Grave).

Otro tanto ocurre con los estadísticos de las SN. En este caso, los promedios obtenidos pusieron de manifiesto la existencia de diferentes tipos de normas percibidas dependiendo de la gravedad de las dificultades de aprendizaje que presentan los alumnos; esto es, los profesores creen percibir de sus

1 Se optó por examinar la actitud de los profesores respecto a esta modalidad por cuanto, conforme a lo expresado por autores como Johnson y Johnson (1984), Molina (1986) o Casanova (1990), aunque presenta mayor número de exigencias a los docentes, es la única que permite alcanzar la normalización.

2 Tales alumnos habían sido seleccionados previamente en orden a representar distintos grados de deficiencias dentro de los cuatro grupos más generales de minusválías.

«otros significativos» prescripciones más favorables a integrar en la medida que los casos sean poco problemáticos.

Por lo que respecta a la relación entre las AB y SN analizadas, los coeficientes de correlación hallados revelaron la existencia de una asociación positiva y significativa entre ambas variables. Lo que nos permite afirmar que los profesores de la muestra, en la medida que son más favorables a integrar a un determinado alumno, tienden a tener al mismo tiempo una normas subjetivas más propensas a dicha acción. No obstante, hemos de señalar que los coeficientes de correlación obtenidos distan de ser perfectos, siendo en su mayor parte moderados.

En lo que se refiere a las relaciones que tanto las AB como las SN mantienen con las BI, los resultados obtenidos revelan que los profesores, en la medida que se muestran más dispuestos a integrar a tiempo completo en sus clases a un determinado alumno, tienden a tener a su vez a tener unas actitudes más favorables y unas normas subjetivas más propensas hacia la realización de dicha acción.

Por último, en lo relativo al papel de las AB y las SN como predictores de las BI, los resultados de los análisis de regresión múltiple efectuados para cada uno de los casos estudiados, revelaron que tanto las actitudes como las normas subjetivas presentaban coeficientes BETA significativos; siendo siempre más elevados los correspondientes a la primera de estas variables. En este sentido, se observó que mientras las AB y las SN explicaban entre un 52-88% de la varianza de las BI, la contribución específica de las normas subjetivas era solo del orden del 1.6-10%. Por lo que podemos afirmar que, si bien las consideraciones valorativas y normativas de los profesores actúan como predictores de las intenciones que estos mantienen respecto a integrar a tiempo completo en sus clases a determinados Alumnos-Problema, es el ámbito actitudinal el que contribuye en mayor medida a la predicción de dichas BI.

DISCUSIÓN

A partir de los resultados obtenidos, parece claro que no se puede hablar de actitudes o de normas homogéneas hacia la integración de «alumnos diferentes», dado que varían en función de la problemática que conllevan las minusvalías de dichos alumnos. Ello es explicable, si tenemos en cuenta el hecho de que para los enseñantes —conforme a lo observado en una investigación previa (Feliciano, 1990)— es fundamental el criterio «grado de adecuación de la deficiencia al estilo de enseñanza de las aulas ordinarias».

Tales diferencias encuentran su correlato en la existencia de intenciones menos favorables hacia el hecho de integrar a un alumno en la medida que éste tenga mayores dificultades para seguir el ritmo de una clase ordinaria. A este respecto, el hecho de que las AB aparezcan como primer predictor de las BI nos hace pensar que si bien las valoraciones personales y las prescripciones percibidas determinan las intenciones, es el dominio de las primeras el que se muestra más relevante para los enseñantes a la hora de establecer su línea de actuación; es más, puede que parte de lo que el profesorado estima como norma percibida sea un reflejo de las propias posturas personales en la medida que consideren que los demás comparten sus puntos de vista, lo cual explicaría la relación constatada entre AB y SN.

REFERENCIAS

- AJZEN, I. y FISHBEIN, M. (1980): *Understanding Attitudes and Predicting Social Behavior*. New Jersey: Prentice Hall, Englewood Cliffs.
- CASANOVA, M. (1990): *Educación Especial: Hacia la integración*. Madrid: Escuela Española.

- FELICIANO, L. (1991): Criterios del profesorado para clasificar alumnos de integración: Aportaciones al escalamiento multidimensional. *Curriculum*, 2, 102-109.
- HORNE, M. (1985): *Attitude toward handicapped students*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- JOHNSON, D. y JOHNSON, R. (1984): Classroom learning structure and attitudes toward students in mainstreaming settings. En R. JONES (Ed.): *Attitude and attitude change in special Education* (pp. 118-142). Virginia: The Council for Exceptional Children.
- MOLINA, S. (1986): La integración de los deficientes mentales. En Fundación del Banco Exterior (Ed.), *Integración en EGB: Una nueva escuela* (pp. 113-132). Madrid: Montalvo.

CREENCIAS CONDUCTUALES DEL PROFESORADO DE EGB SOBRE LOS «ALUMNOS DIFERENTES»: UN ESTUDIO APLICADO A LA INTEGRACIÓN DEL NIÑO ARTRÍTICO REUMATOIDE DESDE LA PERSPECTIVA DE LA TEORÍA DE LA ACCIÓN RAZONADA

Luis Feliciano García y Concepción Riera Quintana
Universidad de La Laguna

INTRODUCCIÓN

De acuerdo con lo expresado por autores como De la Calle (1985) o Giné (1986), no es posible hacer descansar el éxito de un plan de integración solamente en el tesón de algunas personas que intenten ponerlo en marcha y en la modificación de los elementos organizativos del sistema escolar. Junto a ello, es preciso que quienes van a intervenir en el mismo se muestren proclives a cooperar. Ahora bien, **el deseo de participar va a depender, en última instancia, de la existencia de actitudes favorables hacia la integración**; de ahí la importancia que éstas tienen para su implantación y posterior evolución (Molina, 1986).

Conforme a ello, si bien es verdad que los padres de los alumnos y los equipos directivos de los centros desempeñan un papel relevante en el proceso de **normalización**, no es menos cierto que la cristalización del mismo se produce dentro de la clase ordinaria, lo cual confiere un especial significado a la actuación de los profesores del alumno integrado y, por tanto, a las actitudes que estos mantienen al respecto.

Las actitudes que los enseñantes muestran hacia la integración han sido objeto de diferentes estudios (Junkala, 1986; Molina, 1986; Feliciano, 1993), donde se ha puesto de manifiesto que **el profesorado valora de distinta manera la integración de los sujetos con n.e.e en el aula ordinaria dependiendo de la gravedad del déficit que presente**: no estimándose conveniente la incorporación de aquellos alumnos cuyo handicap sea severo, mientras que se considera apropiada la de los sujetos que presentan deficiencias medias o ligeras.

Ahora bien, de acuerdo con lo que plantean Ajzen y Fishbein (1980) en su Teoría de la Acción Razonada (T.A.R.), conocer el tipo de actitudes que los enseñantes tienen en torno al tema no es suficiente para poder *comprender en profundidad* sus intenciones de conducta (**BI**); en este caso, se hace necesario entender por qué el profesorado exhibe dichas actitudes.

Profundizando sobre tales cuestiones, la T.A.R. propone que *el componente actitudinal de las BI es un producto de dos subcomponentes*: a) Las **creencias conductuales salientes** o creencias más

prominentes que tienen los individuos acerca de las posibles consecuencias que se derivan de una acción; b) las **evaluaciones que van asociadas a cada una de esas creencias**. Conforme a ello, Ajzen y Fishbein (1980) señalan que la actitud hacia una conducta (**AB**) puede ser estimada multiplicando la valoración personal (**e**) de cada consecuencia saliente por la probabilidad subjetiva o fuerza de la creencia (**bi**) de que la conducta lleva a la consecuencia en cuestión, y sumando posteriormente los productos obtenidos para la totalidad de las creencias. De este modo, se considera que cuando las personas creen que un determinado comportamiento conduce en general a consecuencias positivas, mantendrán una postura favorable a la realización del mismo, mientras que cuando piensan que conllevan a resultados en su mayoría negativos su actitud será desfavorable. Por tanto, en opinión de los citados autores, si las **AB** tienen un peso significativo en la predicción de las **BI**, el análisis de sus determinantes ofrecerá una estimable ayuda para comprender el proceso por el que las personas configuran determinadas intenciones, y para identificar aquellos elementos sobre los que se ha de intervenir si queremos modificar estas últimas.

Teniendo en cuenta tales planteamientos, y partiendo de los resultados de una investigación previa (Feliciano, 1992) en la que se ponía de manifiesto una alta relación entre las **AB** y las **BI** que los enseñantes exhibían respecto a la integración a tiempo completo de varios Alumnos-Problema, **el estudio que a continuación pasamos a exponer se centró en el análisis de las creencias salientes que el profesorado mantenía acerca de la integración en sus clases a tiempo completo de un determinado sujetos con n.e.e.¹**

MÉTODO

En esta investigación participaron 297 profesores en activo (217 mujeres y 80 varones) pertenecientes a 63 centros de EGB públicos y privados de las islas de Tenerife y Gran Canaria, de los cuales 172 impartían en CS y 125 en PRESC, CI y CM.

Con el fin de estudiar la cuestión anteriormente citada, se elaboró un cuestionario para el profesorado en el que se incluían apartados relativos a: i) datos profesionales y de los centros; ii) la **AB** hacia el hecho de integrar a tiempo completo en sus aulas al Alumno-Problema (la cual era examinada mediante un Diferencial Semántico formado por tres escalas bipolares de adjetivos evaluativos; iii) las (**bi**) relativas a ese hecho (examinadas mediante una escala bipolar de probabilidad); v) las (**e**) de esas creencias conductuales (examinada mediante una escala bipolar de adjetivos evaluativos).

Las 19 creencias conductuales analizadas fueron previamente seleccionadas siguiendo las indicaciones que al respecto formulaban Ajzen y Fishbein (1980): a) Se pasó a una muestra de profesores de EGB un cuestionario abierto en el que se les pedía que enumeraran por separado las consecuencias positivas y negativas que en su opinión se derivarían de la realización de la citada conducta; b) las respuestas positivas y negativas fueron clasificadas por separado en categorías teniendo en cuenta la similitud de las consecuencias a las que aludían; c) se contabilizó el número de veces que había sido citada cada una de las consecuencias incluidas dentro de una categoría, tomando como frecuencia total de esta última la suma de las frecuencias correspondientes a esas consecuencias; d) se estableció que 12, como máximo sería el número de categorías que podían ser consideradas como modales tanto

1 La selección de dicho Alumno-Problema fue llevada a cabo teniendo en cuenta su clasificación en las características que definían a las dimensiones de juicio analizadas en una investigación anterior (Feliciano, 1990). En este caso, optamos por elegir un Alumno-Problema que no estuviese situado en los puestos superiores o inferiores de dichas clasificaciones; de tal modo que, a la hora de considerar su integración a tiempo completo, generase entre los profesores creencias y valoraciones discrepantes, en lugar de reacciones unánimes (a favor o en contra). Conforme a ello, nos pareció adecuado analizar la cuestión aludida en relación al alumno Artrítico Reumatoide; el cual es categorizado por los enseñantes en un nivel medio de dificultad.

dentro de las de carácter positivo como en las de índole negativa; e) se eligió la consecuencia que en cada categoría modal tenía mayor frecuencia absoluta.

El proceso de análisis de datos comprendió: a) El examen de los estadísticos de las variables estudiadas; b) el examen de las relaciones entre las citadas variables mediante el coeficiente de correlación de Pearson y el contraste de medias.

RESULTADOS

Por lo que se refiere a la relación entre las actitudes y el conjunto de creencias relativas a la integración del alumno Artrítico Reumatoide, el coeficiente de correlación hallado evidenció la existencia de una correlación positiva y significativa ($r=0.91$; $p<0.0001$); de lo que se desprende que los profesores de la muestra en la medida que son más favorables a integrar a tiempo completo en sus clases a dicho alumno, tienden a su vez a tener un conjunto de creencias más positivas en torno a dicha acción. A este respecto, los resultados de los contrastes de medias que efectuamos entre cada una de las (bi) y (e) del 25% superior e inferior de los docentes en el baremo de la (AB), pusieron de manifiesto que, si bien existían diferencias significativas entre las (e) de cada subgrupo, los profesores coincidían en el carácter positivo o negativo de dichas valoraciones. Igualmente se comprobó que las consecuencias evaluadas positivamente eran consideradas más probables por los docentes más favorables a la integración de dicho alumno que por los menos favorables, mientras que las evaluadas negativamente eran estimadas más improbables por los primeros que por los segundos; siendo estas diferencias significativas en todos los contrastes.

DISCUSIÓN

Tal y como se pudo comprobar para el caso del alumno Artrítico Reumatoide, en la medida que los profesores difieren en sus AB, también discrepan de las consecuencias que según ellos se derivan de la realización de la conducta en cuestión; advirtiéndose que tales diferencias se daban básicamente en torno a tres tipos de repercusiones relativas a: 1) El ejercicio profesional; 2) el desarrollo social, emocional y cognoscitivo del alumno; 3) el aprendizaje de los demás alumnos y sus reacciones ante el niño integrado. Esas discrepancias eran mayores en las dos últimas categorías (para los Profesores con AB favorables era probable que las consecuencias fueran positivas, siendo improbable para quienes tenían AB desfavorables); mientras que en la primera (aunque en diferente grado) existía cierta tendencia a considerar más probables las secuelas de carácter negativo. En nuestra opinión, estos resultados pueden ser transferibles para los casos de aquellos escolares que, en opinión del profesorado, presentan un grado de dificultad medio o alto. Desde esta perspectiva, creemos que si se quiere contar con una predisposición de los educadores a integrar en sus clases a niños con n.e.e., se hace preciso que estos mantengan unas AB favorables a dicha acción, y para ello es necesario que se muestren convencidos de que de esa conducta no se derivan consecuencias negativas para el trabajo en clase, el aprendizaje de sus alumnos, el clima social del aula o el desarrollo del escolar integrado.

REFERENCIAS

- AJZEN, I. y FISHBEIN, M. (1980): *Understanding Attitudes and Predicting Social Behavior*. New Jersey: Prentice Hall, Englewood Cliffs.
- DE LA CALLE, M. (1985): Aulas especiales: Integración sí pero no. En Actas de la IV Jornadas de E.E. de las EUP de EGB, *La integración del niño disminuido en la escuela ordinaria* (pp. 313-319). Madrid: Ciencias de la Ed. Presc. y Esp.

- FELICIANO, L. (1990): Criterios del profesorado para clasificar alumnos de integración: Aportaciones del escalamiento Multidimensional. *Curriculum*, 2, 102-109.
- FELICIANO, L. (1992): *Análisis estructural de las actitudes: Un estudio aplicado al contexto de la integración escolar*. Tesis Doctoral inédita (F.de Filosofía y Ciencias de la Ed.; Dto. Didáctica e Investigación Ed.; Universidad de la Laguna).
- FELICIANO, L. (1993): Evaluación del proceso de integración: ¿Qué opina el profesorado? En V. ACOSTA (Ed.): *Programas de evaluación e intervención en educación Especial*. (pp. 121-132). La Laguna: Imprecan.
- GINE, C. (1986): La educación especial y la integración de niños disminuidos en Cataluña. En Fundación del Banco Exterior (Ed.): *Integración en EGB: Una nueva escuela* (pp. 37-42). Madrid: Montalvo.
- JUNKALA, J. (1986): Special Education Students in regular classes. *Journal of Learning Disabilities*, 19, 218-221.
- MOLINA, S. (1986): La integración de los deficientes mentales. En Fundación del Banco Exterior (Ed.): *Integración en EGB: Una nueva escuela* (pp. 113-132). Madrid: Montalvo.

EVALUACIÓN DE UN PROGRAMA DE HÁBITOS Y TÉCNICAS DE TRABAJO INTELECTUAL EN LA TRANSICIÓN DE LA EGB AL BUP

*Asunción Ayuso Velar, Andrés Luis Haro, Rosario López Barrionuevo,
Nieves Oreña Castillo, Francisca José Serrano Pastor*
Educadoras/or y Ldas./o. en Pedagogía

1. INTRODUCCIÓN

1.1. Marco teórico

Actualmente existe un gran interés en el desarrollo de cursos de Técnicas de y Hábitos de estudio para los alumnos de todos los niveles educativos; este interés se acentúa sobre todo para aquellos alumnos/as que cursan el ciclo superior de la EGB y las enseñanzas medias (BUP y FP), con el fin de desarrollar los diversos aspectos implicados en el estudio académico (toma de notas, pautas para la realización de trabajos, lectura comprensiva, entre otros), o paliar en la mayoría de los casos las numerosas deficiencias que los estudiantes tienen en los aspectos ya citados. Para Hartley (1986) son tres las razones que han generado la necesidad de desarrollar tales cursos de Técnicas y Hábitos de estudio:

- A) El aumento considerable de la cantidad de información con la que se enfrentan.
- B) Una mayor conciencia de desarrollar habilidades y técnicas de aprendizaje con el fin de manejar eficazmente la información acumulada.
- C) La nueva concepción cognitiva de la psicología del aprendizaje ha propiciado un protagonismo de la persona que como un procesador activo de la información que codifica, interpreta y utiliza en su contexto. Las razones que podríamos ofrecer de por qué el *enseñar a estudiar* cada día está cobrando una mayor relevancia en nuestros contextos educativos son muchas, sin embargo las limitaciones de espacio no nos permiten profundizar en este sentido.

En cualquier caso este interés y necesidad de capacitar al estudiante para que *aprenda a aprender* se refleja en el Libro Blanco para la Reforma del Sistema Educativo (cap. VIII, 1989) en el que se afirma que la capacitación del estudiante consistirá fundamentalmente en la adquisición de las técnicas de base y de las destrezas y estrategias cognitivas para poderse adaptar adecuadamente a sus estudios. La Ley de Ordenación General del Sistema Educativo LOGSE, a través del Diseño Curricular Base, ha supuesto por tanto un cambio cualitativo al menos a nivel legislativo con respecto a la Ley General de Educación de 1970, otorgando a las Técnicas de Trabajo Intelectual (TTI) a través del diseño Curricular Base el estatus de contenidos procedimentales y no ya de objetivos a conseguir.

El presente trabajo es el resultado del interés de la comunidad educativa (padres y madres, profesorado y alumnado) de un colegio privado de Murcia capital por la impartición de cursos de Hábitos y Técnicas de trabajo intelectual a las alumnas del ciclo superior de EGB y de 1º de BUP de dicho centro. La experiencia se viene desarrollando todos los cursos académicos desde que en 1988-89 se iniciara. Los cursos desde su comienzo siempre han sido impartidos por un equipo de pedagogos/as externo al centro. A partir de la experiencia desarrollada en años sucesivos hemos querido evaluar el programa de Técnicas y Hábitos de Trabajo Intelectual en un momento del itinerario educativo que consideramos crucial: el paso de las estudiantes de 8º de EGB a 1º de BUP.

1.2. Descripción del curso de técnicas de estudio

El objetivo general del curso fue el de potenciar en las alumnas las técnicas de trabajo intelectual que tienen adquiridas adecuadamente, modificar las que no utilizan convenientemente, así como estimular la asimilación de aquellas otras encaminadas a favorecer sus procesos individuales de aprendizaje y la efectividad de sus rendimientos académicos. Para ello, el programa está dirigido a las alumnas como artífices de sus procesos de aprendizaje, sin desestimar la importancia del resto de implicados en dicho proceso.

En resumen, el contenido del Programa se estructuró en torno a tres grandes ámbitos:

- 1.— Las estrategias y procedimientos necesarios para recoger la información para aprender.
- 2.— La motivación, elemento fundamental en todo proceso de aprendizaje.
- 3.— Habilidades y destrezas específicas que la alumna debe poseer, conocer y utilizar en su tarea escolar.

Estos bloques temáticos se impartieron durante ocho sesiones de una hora de duración, mediante una metodología con dos vertientes diferenciadas aunque complementarias: por un lado, se utilizó por parte del equipo orientador una metodología expositiva para poner en conocimiento de las alumnas los contenidos básicos del programa que se estaba poniendo en práctica; por otro lado, también se recurrió a una metodología activa-participativa, con el fin de poner en práctica, mediante los materiales que posteriormente señalaremos, los contenidos del programa expuestos a las alumnas. Éstas trabajaban en pequeños grupos bajo la supervisión del equipo orientador. De esta manera, la formación de las alumnas comprendía tres aspectos básicos: la adquisición de contenidos básicos, su puesta en práctica supervisada por el equipo orientador y el trabajo cooperativo.

Otro aspecto importante en la aplicación del programa se refiere al control de las variables extrañas que podrían afectar a la investigación, teniendo en cuenta que ésta abarca dos cursos escolares. Dicho control se ha visto facilitado por el hecho de que la intervención ha sido realizada por el mismo equipo de personas en las dos ocasiones.

Esta situación hace que la investigación goce de unas condiciones de aplicación y control del tratamiento altamente fiables, considerando la posible variabilidad.

En la aplicación del programa hay otro aspecto a tener en cuenta: las alumnas trabajaban los contenidos con un material específico. Este material, junto con el utilizado para elaborar el programa, es de suma importancia para todos los aspectos de la investigación llevada a cabo; el material utilizado fue: material elaborado por el equipo orientador que realizó el programa para trabajar los distintos contenidos del curso; y material adaptado de autores/as como Hernández Pina (1990), Álvarez y Fernández (1990) y Carlos Gómez, García y Alonso (1991).

2. OBJETIVOS

Nuestra pregunta de investigación se ciñe al interrogante siguiente: *¿mejoran los hábitos y técni-*

cas de trabajo intelectual de las alumnas mediante un curso intensivo de TTI impartido por un equipo orientador externo al centro, en la transición de la EGB al BUP?

Intentaremos dar respuesta a través de la formulación de los objetivos que sintetizamos a continuación:

1º. Conocer la situación inicial de las distintas dimensiones de hábitos y técnicas de trabajo intelectual de las alumnas (en 8º de EGB y antes de iniciar el curso de TTI).

2º. Determinar la situación final de las distintas dimensiones de hábitos y técnicas de trabajo intelectual de las estudiantes al inicio del siguiente curso académico (ya en 1º de BUP).

3º. Analizar las mejoras en los distintos aspectos de hábitos y técnicas de trabajo intelectual evaluados que se han producido en las alumnas, en la transición de 8º de EGB a 1º BUP.

3. METODOLOGÍA

3.1. Sujetos

La experiencia se desarrolló con las alumnas del citado centro privado matriculadas oficialmente en los cursos académicos 1990-91 y 1991-92, durante el cual cursaban 8º de EGB, y aquéllas que pasaron a 1º de BUP los cursos 1991-92 y 1992-93 respectivamente.

Para el desarrollo del curso de TTI, así como para la aplicación del instrumento de evaluación se respetaron las aulas intactas. Además, cabe señalar que la cumplimentación del cuestionario fue realizada por aquellas alumnas que asistieron a la sesión en la que estaba diseñada esta actividad, aunque no tuvieran conocimiento previo de la misma y accedieran voluntariamente a ello.

Las muestras totales para ambos grupos se muestran en la tabla siguiente:

	Frecuencias	Porcentajes
Cursos 1990-91/1991-92	79.00	42.70
Cursos 1991-92/1992-93	106.00	57.30
Total	185.00	100.00

3.2. Recogida de datos

Instrumento

El instrumento de recogida de datos ha sido el *Cuestionario de Hábitos y Técnicas de Estudio* (CHTE) de Álvarez y Fernández (1990). Su finalidad es diagnosticar a nivel individual y grupal los hábitos y técnicas de estudio para plantear un programa de intervención en el aula, que contemple todos aquellos aspectos que sean susceptibles de mejora.

En la elaboración del CHTE se han considerado las condiciones físicas y ambientales, la planificación y estructuración del tiempo y el conocimiento de las técnicas básicas. Estos aspectos se han desglosado en las siguientes escalas:

AC — Actitud general hacia el estudio, incluye todo lo que hace referencia a la predisposición, interés motivación hacia el estudio.

LU — Lugar de estudio, se refiere no sólo al sitio físico sino también a las características del entorno que inciden en la concentración y rendimiento.

ES — Estado físico del escolar, se refiere a las condiciones físicas personales que le permitan un buen rendimiento en el estudio.

PL — Plan de trabajo, incluye todo lo que hace referencia a una buena planificación y estructuración del tiempo que se va a dedicar al estudio, teniendo en cuenta el número de materias y su dificultad.

TE — Técnicas de estudio, ofrece pautas de como estudiar y recoge los diferentes pasos que debe seguirse para el estudio de un tema.

EX — Exámenes y ejercicios, se refiere a las pautas que conviene seguir cuando se va a realizar un examen o ejercicio.

TR — Trabajos, incluye aspectos que se han de tener en cuenta para realizar un trabajo.

Cada escala comprende una serie de elementos y entre las siete reúnen un total de 56. Los materiales utilizados para la administración y cumplimentación del cuestionario fueron: manual del CHTE, cuadernillo de preguntas para cada alumna y hoja de corrección y perfil para cada alumna.

Procedimiento

La recogida de datos se realizó en la primera sesión de los cursos, de una hora de duración para cada uno de los grupos, en horario lectivo.

Las alumnas no fueron avisadas previamente de la administración del cuestionario. Fue en esta primera sesión cuando se les informó de la finalidad de la misma y se les solicitó su colaboración voluntaria.

Para la aplicación se procedió de la forma siguiente:

- Se explicó a las alumnas las instrucciones de cumplimentación del cuestionario.
- Se entregó a cada estudiante el cuestionario al que debían responder.
- Durante la aplicación se resolvieron todas las dudas planteadas por las alumnas.

Procedimiento para la corrección del cuestionario

Tras la cumplimentación del cuestionario cada alumna procedió a la corrección del mismo, y a la elaboración de su perfil individual de Hábitos y Técnicas de Estudio.

La corrección y los cálculos necesarios para ello fueron realizados por las propias alumnas bajo las instrucciones y la supervisión constantes del equipo orientador.

En concreto, el procedimiento de corrección se realizó del siguiente modo:

1º. Las alumnas marcaron en la hoja de corrección que se les facilitó las respuestas al cuestionario, de acuerdo con el modelo normativo de corrección que marca el CHTE obteniendo las puntuaciones directas en cada una de las escalas.

2º. A partir de las puntuaciones directas las estudiantes tomaron nota de los percentiles que le correspondían en cada una de las escalas, según los listados que el equipo orientador les facilitó correspondientes al baremo adecuado a su nivel educativo.

3º. Aunque estas puntuaciones se dieron como correctas para la elaboración del perfil individual de Hábitos y Técnicas de Estudio, con el fin de cubrir los objetivos propuestos para esa sesión didáctica —que las alumnas tomaran conciencia de cuál era su situación de los Hábitos y Técnicas de Estudio—, tanto las puntuaciones directas como su transformación en percentiles fueron revisadas posteriormente por el equipo de trabajo para la elaboración de los informes destinados al profesorado, tutores/as y para su utilización en la presente investigación.

3.3. Variables de estudio

Las variables dependientes de nuestro estudio son en total 14, las cuales se corresponden con las puntuaciones directas obtenidas en seis escalas de las siete que incluye el CHTE (hemos excluido la escala ES, estado físico de la alumna) y para cada uno de los cursos académicos en los que se ha aplicado el mencionado cuestionario (variables independientes). Además, hemos contemplado dos variables dependientes más, que se corresponden con el sumatorio de las seis puntuaciones directas del cuestionario de cada uno de los años académicos. En concreto las variables dependientes son las siguientes:

	PRE-TEST	POST-TEST
— Actitud general hacia el estudio	AC	ACPOST
— Lugar de estudio	LU	LUPOST
— Planificación del estudio	PL	PLPOST
— Técnicas de Estudio	TE	TEPOST
— Exámenes	EX	EXPOST
— Trabajos	TR	TRPOST
— Total	TOTAL	TOPOST

3.4. Plan de análisis de datos

Inicialmente realizamos un análisis exploratorio de los datos para identificar los valores fuera de rango, las puntuaciones extremas y las anomalías en la forma y simetría de las distribuciones. Para ello construimos un gráfico de caja (*box plot*), histogramas digitales (*stem and leaf*) y gráficos de probabilidad (*normal probability plot*). A continuación llevamos a cabo una lectura directa de las variables de estudio y por último, a pesar de haber depurado nuestros datos como éstos continuaban sin mantener las condiciones necesarias para la aplicación de pruebas estadístico-paramétricas (Ato y otros, 1990), optamos por la aplicación de una prueba estadística no paramétrica para el diseño de dos muestras relacionadas: la prueba de rangos de Wilcoxon (más precisa que la prueba de signos) por tratarse nuestros datos de datos no binarios o dicotómicos y que, por tanto, tiene en cuenta el rango u orden de cada diferencia (pensemos que la prueba de signos sólo considera la dirección positiva o negativa de las diferencias).

4. ANÁLISIS ESTADÍSTICOS Y RESULTADOS

Como resultado del análisis exploratorio se han eliminado 19 sujetos de los 79 de la muestra inicial del curso 1990-91 y 1991-92 que presentaban puntuaciones extremas, y 27 de los 106 sujetos de la muestra inicial de los cursos 1991-92 y 1992-93; la muestra definitiva pues paso a 139 alumnas en total. Los resultados obtenidos con la lectura directa de las variables dependientes para los grupos 1990-91/1991-92 y 1991-92/1992-93 respectivamente son los siguientes:

GRUPO 1990-91/1991-92

	Nº de casos	Valor mínimo	Valor máximo	Media	Desviación típica
AC	60	5.000	10.000	8.217	1.606
ACPOST	60	5.000	10.000	8.333	1.398
LU	60	5.000	10.000	8.933	1.287
LUPOST	60	8.000	10.000	9.483	0.701
PL	60	3.000	10.000	5.817	1.882
PLPOST	60	1.000	10.000	6.200	2.223
TE	60	3.000	9.000	6.417	1.476
TEPOST	60	2.000	9.000	6.750	1.865
EX	60	2.000	5.000	3.900	0.796
EXPOST	60	1.000	5.000	4.083	0.944
TR	60	2.000	6.000	4.017	1.000
TRPOST	60	3.000	6.000	4.467	0.911
TOTAL	60	31.000	53.000	41.733	5.204
TOPOST	60	29.000	55.000	43.467	6.39

GRUPO 1991-92/1992-93

	Nº de casos	Valor mínimo	Valor máximo	Media	Desviación típica
AC	79	5.000	10.000	8.329	1.421
ACPOST	79	4.000	10.000	8.380	1.496
LU	79	5.000	10.000	9.051	1.176
LUPOST	79	8.000	10.000	9.481	0.714
PL	79	1.000	10.000	6.722	1.928
PLPOST	79	1.000	10.000	6.430	2.158
TE	79	2.000	9.000	5.405	1.822
TEPOST	79	3.000	9.000	6.506	1.543
EX	79	3.000	5.000	4.215	0.673
EXPOST	79	2.000	5.000	3.835	0.898
TR	79	2.000	5.000	3.595	0.954
TRPOST	79	2.000	5.000	3.620	0.991
TOTAL	79	30.000	53.000	41.316	5.850
TOPOST	79	26.000	54.000	42.747	5.326

Por otro lado los resultados que hemos obtenido con la lectura directa de tales variables para la muestra total definitiva son los que presentamos a continuación:

	Nº de casos	Valor mínimo	Valor máximo	Media	Desviación típica
AC	185	3.000	10.000	8.135	1.549
ACPOST	185	4.000	10.000	8.211	1.548
LU	185	4.000	10.000	8.859	1.315
LUPOST	185	4.000	10.000	9.135	1.160
PL	185	0.000	10.000	6.081	2.182
PLPOST	185	1.000	10.000	6.157	2.283
TE	185	1.000	9.000	5.616	1.859
TEPOST	185	0.000	9.000	6.476	1.742
EX	185	0.000	5.000	3.908	0.907
EXPOST	185	0.000	5.000	3.827	0.996
TR	185	0.000	6.000	3.659	1.087
TRPOST	185	0.000	6.000	3.805	1.196
TOTAL	185	22.000	53.000	40.335	6.278
TOPOST	185	26.000	55.000	41.832	6.502

Como ya hemos apuntado anteriormente las distribuciones de las variables de nuestro estudio no se ajustaban a las condiciones de normalidad en ningún caso, por lo que tuvimos que optar por la aplicación de una prueba estadística no paramétrica. Los resultados obtenidos al aplicar la prueba de rango de Wilcoxon y que resultan estadísticamente significativos fueron los siguientes:

GRUPO 1990-91/1991-92	Probabilidad
LU-LUPOST	0.001
TR-TRPOST	0.010
TOTAL-TOPOST	0.021

GRUPO 1991-92/1992-93	Probabilidad
LU-LUPOST	0.005
TE-TEPOST	0.000
EX-EXPOST ***	0.004
TOTAL-TOPOST	0.038

TOTAL MUESTRA DEFINITIVA	Probabilidad
LU-LUPOST	0.000
TE-TEPOST	0.000
TOTAL-TOPOST	0.002

5. CONCLUSIONES

A la vista de los resultados podemos extraer las siguientes conclusiones:

- Podemos afirmar que en ambos grupos se produce una mejora global en los distintos aspectos de hábitos y técnicas de trabajo intelectual evaluados, si bien en el grupo (91-92/92-93) no se ha producido mejora en la planificación personal del estudio ni en las pautas para la realización de los exámenes. A pesar de ello, tanto en el total de la muestra total definitiva como en los totales de las muestras parciales de cada uno de los grupos, se observan mejoras estadísticamente significativas.

- Respecto a los aspectos específicos de TTI, en los que se observan diferencias estadísticamente significativas que impliquen una mejora de las mismas tenemos:

- Para el total de la muestra, el lugar y las técnicas de estudio.

- Para el grupo (90-91/91-92), el lugar de estudio y las pautas para la realización de trabajos escritos.

- Para el grupo (91-92/92-93), el lugar y las técnicas de estudio.

En consecuencia, las condiciones del lugar de estudio mejoran significativamente en los dos grupos, así como en la muestra total. Los otros dos aspectos —las técnicas de estudio y las pautas para la realización de trabajos escritos—, si bien también para el primer grupo, solamente en el segundo grupo y el total las diferencias son estadísticamente significativas.

- En el segundo grupo, observamos que tanto en la planificación personal del estudio como en las pautas para la preparación y realización de exámenes, no sólo no han mejorado sino que además se produce un empeoramiento. Es más, en el caso de las pautas para la preparación y realización de exámenes, la diferencia obtenida es estadísticamente significativa. Este último resultado vuelve a ratificarse en la muestra total, aunque en esta ocasión no de forma estadísticamente significativa.

- A la vista de los resultados también señalar que, a nuestro juicio, parece que son aquellos aspectos más externos a las alumnas (relacionados con las tareas) los más susceptibles de ser adquiridos y/o modificados (condiciones del lugar de estudio, las pautas para la realización de trabajos escritos y las técnicas de estudio). Por el contrario, aquellos aspectos más internos (actitud hacia el estudio, planificación personal y pautas para la preparación y realización de exámenes) son más difíciles de adquirir y/o modificar mediante la instrucción.

Pensamos, así mismo, que los cursos TTI impartidos por un equipo orientador externo y de forma intensiva, en general producen una mejora en los hábitos y técnicas de trabajo intelectual de las alumnas (así se ha puesto de manifiesto). Probablemente estas mejoras serían superiores si los cursos de TTI tuvieran un planteamiento curricular y se prolongaran durante todo el curso académico (por no decir durante todo el itinerario académico del alumno/a), teniendo en cuenta los contenidos y la naturaleza de la información que específicamente demanda el planteamiento curricular. En cualquier caso, las intervenciones en este sentido deberán profundizar más en los aspectos internos de las TTI sin olvidar evidentemente los más externos.

BIBLIOGRAFÍA

ÁLVAREZ, M. y FERNÁNDEZ, R. (1990): *Cuestionario de Hábitos y Técnicas de Estudio. Manual*. Madrid: TEA Ediciones, S.A.

ATO, M., LÓPEZ, J. A., VELANDRINO, A. P. y SÁNCHEZ, J. (1990): *Estadística avanzada con el paquete Systat*. Murcia: Universidad de Murcia.

CARLOS, P., GARCÍA, A. y ALONSO, P. (1991): *Manual de TTI. Procedimientos para aprender a aprender*. Madrid: EOS.

HARTLEY, J. (1986): *Improving study skills*. British Educational Research Journal 12 (2), 111-126.

HERNÁNDEZ PINA, F. (1990): *Aprendiendo a aprender*. Murcia: la autora.

INTERVENCIÓN PSICOPEDAGÓGICA FAMILIAR: ADAPTACIÓN, APLICACIÓN Y EVALUACIÓN DEL PROGRAMA S.T.E.P./TEEN DE DINKMEYER Y MCKAY

I. Bartau & J. M. Maganto
Universidad del País Vasco

Esta comunicación es el resultado de una intervención psicopedagógica familiar. El objetivo general de la misma es adaptar, aplicar y evaluar el Programa S.T.E.P./Teen (Systematic Training for Effective Parenting/adolescencia) de Dinkmeyer y McKay (1990). Asimismo, se pretende analizar la percepción de las prácticas educativas parentales percibidas por el grupo de padres y por sus hijos adolescentes, así como el grado de eficacia y satisfacción percibido por los padres a lo largo del programa.

SUJETOS

En la investigación participaron un total de 70 sujetos, 35 padres y sus 35 hijos adolescentes. El grupo de padres estaba compuesto de 9 hombres y 26 mujeres cuyas edades oscilaban entre 32 y 51 años, pertenecientes a un nivel educativo y socioeconómico medio y únicamente 2 sujetos formaban familias monoparentales. El grupo de hijos, con edades comprendidas entre 14 y 17, en su mayoría cursaban B.U.P./F.P. en centros públicos.

OBJETIVOS

1º) Comparar los estilos de socialización percibidos en el hogar, tanto por los padres como por los hijos, antes y después de la implementación del programa S.T.E.P./teen.

2º) Analizar las actitudes educativas de los padres, antes y después de la implementación del programa S.T.E.P./teen.

3º) Comprobar si el programa S.T.E.P./teen fomenta el aprendizaje de las estrategias propuestas por el programa: comprensión y reacción ante el comportamiento inadecuado de los hijos, procedimientos de comunicación y tipo de ambiente familiar.

4º) Analizar el grado de implicación, eficacia y satisfacción percibido por los padres a lo largo de las sesiones y una vez finalizado el programa.

INSTRUMENTOS DE MEDIDA

— «Estilos de socialización en la familia, (EMBU)» de Musitu y otros (1991) que evalúa seis factores: sobreprotección, comprensión y apoyo, castigo, presión hacia el logro, rechazo y atribución de culpa. Versión padres y versión hijos.

— «Actitudes Educativas de los Padres, (AEP)» de Pérez Sánchez y otro (1991) que mide cuatro factores bipolares: aceptación-rechazo, comunicación-incomunicación, autonomía-sobreprotección y democracia-autoritarismo.

— «Evaluación de las estrategias de comunicación propuestas por el programa S.T.E.P./teen», Bartau y Maganto (1993).

— Autoevaluación de la eficacia y satisfacción percibida por los padres en el programa.

— Autoinforme de evaluación continua acerca del grado de implicación y satisfacción personal de los padres en el programa.

Procedimiento

El programa S.T.E.P./teen pretende estimular la reflexión acerca de las prácticas educativas parentales y ofrece estrategias para mejorar la comunicación padres-hijos adolescentes. Está dirigido a padres que no presentan problemáticas cotidianas especiales. Su metodología requiere la participación activa y continua de los padres y la coordinación de un monitor. Consta de 10 sesiones de dos horas de duración semanales. Se utilizó un diseño antes y después como una primera aproximación a la evaluación de este programa.

Análisis de los resultados

Teniendo en cuenta los objetivos propuestos, podemos hacer las siguientes consideraciones:

1º) En relación a los estilos de socialización familiar percibidos en el hogar, hemos obtenido los siguientes resultados:

— Los estilos de socialización familiar percibidos por padres e hijos no se diferencian antes ni después del programa de intervención.

— Sin embargo, se constatan diferencias en el grupo de padres expuestos al programa de intervención. Por un lado, disminuyen significativamente los niveles de sobreprotección, castigo y presión hacia el logro, y, por otro, aumenta el grado de comprensión y apoyo percibido por los padres después de su exposición al programa. Por último, no se observan diferencias en relación al rechazo y atribución de culpa.

— Tampoco se aprecian diferencias con respecto a los estilos de socialización percibidos por los hijos antes y después de exponer a sus padres al programa.

2º) Del análisis de la evaluación de las actitudes educativas de los padres antes y después del programa, señalar que aumenta significativamente el nivel de aceptación de los padres con respecto a los hijos, disminuye el grado de sobreprotección y autoritarismo, y no se observan diferencias en el resto de los factores actitudinales considerados.

3º) Con respecto a la evaluación del aprendizaje de las estrategias propuestas por el programa, hemos obtenido los siguientes resultados:

— Se constatan que, una vez finalizado el programa, los padres aumentan significativamente su capacidad de comprensión y reacción positiva ante el comportamiento inadecuado de sus hijos. En este sentido, aumenta la aceptación de los hijos y la capacidad de los padres de control ante las exigencias de atención, enfrentamientos de poder y demostración de insuficiencia por parte de los hijos.

ESTILOS DE SOCIALIZACIÓN: PERCEPCIÓN PADRES.
COMPARACIÓN DE RESULTADOS ANTES Y DESPUÉS DEL PROGRAMA

VARIABLE	N	X	S	P
SOBREPROTECCIÓN	14	19.35	1.82	.002
		16.85	1.99	
COMPRESIÓN Y APOYO	12	51.33	5.64	.069
		54.41	4.92	
CASTIGO	13	20.69	1.75	.001
		18.53	1.26	
PRESIÓN HACIA EL LOGRO	14	20.78	3.09	.003
		18.35	1.98	
RECHAZO	12	12.00	2.04	.204
		11.08	1.50	
ATRIBUCIÓN DE CULPA	12	20.50	2.46	.105
		19.08	1.16	

— Por otra parte, hemos encontrado diferencias en cuanto al tipo de procedimientos de comunicación que utilizan los padres para relacionarse con sus hijos después de participar en el programa. Concretamente, aumenta significativamente la atención reflexiva, mensajes-yo, consecuencias lógicas, aceptación de la responsabilidad, estimulación y autoaceptación personal.

— Por último, con respecto al tipo de ambiente familiar, señalar diferencias significativas en cuanto al ambiente autónomo, en el sentido de que los padres, después de finalizado el programa, permiten en mayor medida que sus hijos tomen decisiones y se responsabilicen de ellas.

4º) Respecto a la evaluación continua de las sesiones, señalar, que si bien sólo la mitad de los padres realizaban las actividades prácticas y ejercicios propuestos para el hogar, la mayoría acudía a las sesiones habiendo leído el capítulo correspondiente a cada sesión. Los padres consideran que pudieron resolver las dudas que se les planteaban y comprender así mejor los conceptos y estrategias que ofrece el programa, percibiéndolas además como novedosas y útiles.

Por otro lado, la mayoría de los padres presentan un nivel de participación y satisfacción muy elevado a lo largo de las sesiones.

En la siguiente tabla podemos observar que, antes de comenzar el programa, los padres presentaban un alto grado de motivación e interés por el tema que se mantuvo a lo largo del mismo.

ACTITUDES EDUCATIVAS DE LOS PADRES.
COMPARACIÓN DE RESULTADOS ANTES Y DESPUÉS DEL PROGRAMA

VARIABLE	N	X	S	P
ACEPTACIÓN	14	16.00	2.66	.008
		18.14	1.95	
RECHAZO	13	10.30	5.02	.823
		10.00	5.98	
COMUNICACIÓN	15	35.60	4.53	.876
		35.46	3.72	
INCOMUNICACIÓN	13	8.00	5.80	.900
		7.76	6.87	
AUTONOMÍA	11	27.90	4.23	.991
		27.00	3.28	
SOBREPROTECCIÓN	14	15.07	3.75	.032
		11.14	6.78	
DEMOCRACIA	14	29.42	4.07	.332
		30.64	4.97	
AUTORITARISMO	13	12.76	5.76	.009
		8.38	6.09	

En cuanto al grado de eficacia y satisfacción percibido por los padres una vez finalizado, consideran que este programa estimula la reflexión acerca de la prácticas educativas parentales en el hogar y que promueve el aprendizaje de nuevas habilidades y estrategias para comprender y reaccionar ante el comportamiento de sus hijos más satisfactoriamente.

Con respecto a los procedimientos de comunicación que ofrece el programa para mejorar las relaciones padres-hijos, los padres expresan haberlos utilizado en la práctica cotidiana, a excepción de la reunión familiar que apenas la han puesto en práctica. Asimismo, piensan que son muy eficaces y

ESTRATEGIAS DE COMUNICACIÓN PADRES-HIJOS (STEP-TEEN):
CONTROL DEL COMPORTAMIENTO ANTES-DESPUÉS.

VARIABLE	N	X	S	P
CONTROL ATENCIÓN	14	3.07	1.73	.012
		5.28	2.23	
CONTROL PODER	13	6.53	1.80	.002
		8.84	1.86	
CONTROL INSUFICIENCIA	14	7.35	2.06	.007
		8.85	1.56	
CONTROL REVANCHA	13	7.15	1.72	.120
		8.38	1.85	

que esperan utilizarlos más asiduamente en el futuro. También valoran positivamente los materiales utilizados y la coordinación de las monitoras.

Es importante señalar que los padres valoran el programa como altamente recomendable para otros padres de familia con características similares.

Estos resultados se confirman en las respuestas abiertas formuladas en el cuestionario y en los autoinformes diarios de las sesiones.

CONCLUSIONES

En primer lugar, apreciamos cambios significativos tanto en la aproximación a la evaluación externa como interna del programa, que exponemos a continuación.

Del análisis de los resultados, podemos concluir que aunque este programa de intervención familiar no repercute en los estilos de socialización familiar percibidos por los hijos, incide significativamente en los estilos de socialización y actitudes educativas de los padres. Concretamente, favoreciendo la capacidad de aceptación, comprensión y apoyo de los hijos y disminuyendo los niveles de sobreprotección, autoritarismo, castigo y presión hacia el logro. Por lo tanto, creemos que este programa contribuye a que los padres desarrollen prácticas educativas más democráticas que favorezcan la autonomía y aceptación de los hijos.

Por otra parte, se produce un cambio significativo en la percepción de los padres en torno a las estrategias de comunicación utilizadas con sus hijos. Después de la intervención, los padres se sienten más capaces de escuchar y comprender a sus hijos, de ponerse en su lugar y de expresarles adecuadamente sus sentimientos. Además creen que pueden fomentar la responsabilidad y autoestima de sus hijos, estimulándoles y haciéndoles asumir las consecuencias de su comportamiento.

ESTRATEGIAS DE COMUNICACIÓN PADRES-HIJOS (STEP-TEEN):
PROCEDIMIENTOS ESPECÍFICOS ANTES-DESPUÉS

VARIABLE	N	X	S	P
ATENCIÓN REFLEXIVA	14	3.07	1.73	.012
		5.28	2.23	
MENSAJES-YO	14	8.00	1.41	.001
		9.28	1.26	
EXPLORACIÓN ALTERNATIVAS	13	7.00	1.22	.211
		7.69	1.97	
CONSECUENCIAS LÓGICAS	13	7.23	1.69	.003
		8.92	1.32	
REUNIÓN FAMILIAR	13	7.23	2.24	.544
		7.61	1.93	
RESPONSABILIDAD CAMBIO	14	8.57	1.50	.023
		9.71	.61	
ACEPTACIÓN	14	16.00	2.66	.008
		18.14	1.95	
AUTOACEPTACIÓN	14	6.50	2.27	.009
		8.42	2.02	
ESTIMULACIÓN	14	26.85	4.41	.001
		32.71	4.08	
PROPIEDAD DEL PROBLEMA	15	12.76	5.76	.009
		8.38	6.09	

ESTRATEGIAS DE COMUNICACIÓN PADRES-HIJOS (STEP-TEEN):
 AMBIENTE FAMILIAR ANTES-DESPUÉS

VARIABLE	N	X	S	P
AMBIENTE CONTROLADO (PERMISIVO)	13	6.38	2.36	.070
		7.92	1.97	
AMBIENTE AUTÓNOMO (SOBREPROTECCIÓN)	14	7.14	1.61	.005
		8.92	1.85	
AMBIENTE DEMOCRÁTICO (AUTOCRÁTICO)	14	7.57	1.55	.701
		7.35	2.49	

AUTOINFORME DE LAS SESIONES (EVALUACIÓN CONTINUA)
 MEDIA DE PORCENTAJES DE LOS RESULTADOS DE LAS DIEZ SESIONES

VARIABLES	% SÍ	% NO	N
PRÁCTICA ACTIVIDAD HOGAR	51	49	17
LECTURA CAPÍTULO HOGAR	85	15	20
REALIZACIÓN EJERCICIOS HOGAR	46	54	16
COMPRENSIÓN CONCEPTOS SESIÓN	100	0	19
PARTICIPACIÓN SESIÓN	85	15	20
NOVEDAD DE APRENDIDO SESIÓN	95	5	19
UTILIDAD ESTRATEGIAS SESIÓN	100	0	19
RESOLUCIÓN DUDAS MONITOR	100	0	21
FOMENTO PARTICIPACIÓN MONITOR	95	5	19
SATISFACCIÓN SESIÓN	95	5	19

Los padres, a nivel personal, confían más en su capacidad de autocontrol ante las situaciones conflictivas cotidianas que se plantean en el hogar. En la resolución de los conflictos adoptan procedimientos más democráticos fomentando el respeto y la colaboración de todos los miembros de la familia.

Otra de las conclusiones que queremos destacar es el alto grado de participación y satisfacción de los padres con respecto al desarrollo del programa. En este sentido, valoraron positivamente este tipo

AUTOINFORME DE EVALUACIÓN FINAL (STEP/TEEN) DESPUÉS DEL PROGRAMA
(1: MUY POCO. 2: POCO. 3: REGULAR. 4: BASTANTE. 5: MUCHO)

VARIABLES	% SÍ	% NO	N
MOTIVACIÓN POR EL PROGRAMA (ANTES)	16	4.06	0.87
MOTIVACIÓN POR EL PROGRAMA DESPUÉS)	16	4.87	0.78
ESTIMULACIÓN DE REFLEXIÓN	16	4.81	0.80
COMPRENSIÓN COMPORTAMIENTO DE HIJOS	16	4.59	1.01
UTILIZACIÓN ESTRATEGIAS	16	3.28	0.78
EFICACIA ESTRATEGIAS	16	4.84	0.83
EXPECTATIVA DE UTILIZACIÓN DE ESTRATEGIAS	16	4.30	0.77
SATISFACCIÓN SESIONES	16	4.87	0.81
PARTICIPACIÓN SESIONES	16	4.50	0.75
EVALUACIÓN MONITORES	16	4.95	0.55
RECOMENDACIÓN PROGRAMA	16	4.90	0.72
UTILIDAD MATERIAL	16	4.80	0.77

de intervención en relación a la experiencia que tenían con otras actividades de carácter más informativo o puntual. Resaltar, además, su interés por continuar este tipo de formación en el futuro.

También hemos observado a lo largo de la intervención, que la reflexión personal y compartida con el grupo que estimula el programa en los padres, les crea inicialmente «conflictos» entre sus propias prácticas educativas y relacionales y las que propone el programa, suscitando un constante replanteamiento entre las mismas, que progresivamente van solucionando según van asimilando e incorporando las distintas alternativas y posibilidades que el propio proceso del programa les brinda.

Por último, destacar que los padres expresaban la necesidad de que sus hijos aprendieran procedimientos de comunicación similares con el fin de contribuir tanto a su propio desarrollo personal como al clima interaccional dentro de la familia. En este sentido, pensando en futuras investigaciones, nos parece interesante desarrollar programas paralelos y complementarios padres-hijos.

EVALUACIÓN DEL PRACTICUM DE «ORIENTACIÓN EDUCATIVA», A PARTIR DE LAS OPINIONES DE LOS ESTUDIANTES Y DE LOS CENTROS DE FORMACIÓN

Alfredo Pérez Boullosa, Ignacio Alfaro Rocher,
Universitat de València
Miguel Ángel Barberá, Laura Ramírez
Fundación Universidad-Empresa de Valencia, ADEIT

INTRODUCCIÓN

En la evaluación del programa de prácticas que ofrece el *departamento M.I.D.E.* de la *Universitat de València* para la formación del Orientador, podemos diferenciar al menos dos objetivos: uno está relacionado con el impacto de esta experiencia en los estudiantes y en los centros educativos dónde la realizan, con una finalidad de mejora de los aspectos estructurales y organizativos y, otro, con la evaluación de las funciones, contenido y objetivos formativos de las mismas. Evidentemente, ambos están relacionados. Sin embargo, este trabajo se centra en el primer objetivo, ya que la iniciativa de esta evaluación y su materialización corresponde a un agente externo: la *Fundación Universidad Empresa Valencia, ADEIT*, promovida por el Consejo Social de la U.V. Por lo que respecta al segundo objetivo, el departamento MIDE viene desarrollando una evaluación continua del programa perfilando y analizando cualitativamente los objetivos y contenidos con cada uno de los centros educativos integrados en dicho programa. Tales especificaciones vienen recogidas en el documento *Programa de Prácticas*, (MIDE, 1992) que anualmente este Departamento edita y distribuye gratuitamente a todos los estudiantes y centros de formación. En cualquier caso, téngase en cuenta que en la formación del orientador existen multiplicidad de técnicas, áreas de competencias a desarrollar, etc. (Echeverría, 1993; Repetto, 1993) y que este programa de prácticas no se estructura conforme a un diseño basado en experiencias de laboratorio, sino que se fundamenta en experiencias reales en centros educativos bajo la supervisión y control del tutor de la empresa (*orientador*) y de la Universidad (*profesor*). Asimismo, la evaluación del estudiante no la realiza fundamentalmente el orientador del centro de acogida como ocurre en otros contextos (Goodman, 1993), sino que los dos profesionales mencionados son los responsables de la calificación del estudiante, que se efectúa de una forma precisa y exhaustiva a través de tres instrumentos y que tiene una representación de un 25% en la calificación de determinadas «asignaturas» (MIDE, 1992).

Por otra parte, el programa de prácticas mencionado se articula a través de la *Fundación Universidad Empresa de Valencia, ADEIT*, la cual desarrolla de forma general un programa de «Prácticas en Empresa», al que pueden acogerse los estudiantes de los dos últimos años de estudios universitarios.

Su tarea se circunscribe al establecimiento de una vinculación entre los «centros de trabajo» y la «universidad», sin participar en absoluto en el diseño y contenido de las prácticas. En este contexto, nuestro programa formaliza cada año unos *Convenios de Cooperación Educativa* individualizados entre cada Centro Docente y la Universitat de València, responsabilizándose de la tutorización del alumno un profesor del Departamento MIDE de la subárea de Orientación y Diagnóstico. El hecho de integrar nuestras prácticas en el programa del ADEIT ofrece innegables ventajas tanto para el alumno (ayudas económicas, certificados, etc.) como para los centros educativos, orientadores de los mismos y, también, para el propio departamento, ya que no sólo institucionaliza este practicum con unas ciertas garantías legales (seguros de responsabilidad civil, de desplazamientos, etc.) sino que ofrece todos sus servicios técnicos para poder realizar los trámites administrativos y, como es el caso de esta comunicación, supervisar como evaluador externo su gestión del prácticum.

OBJETIVOS

Los datos que se presentan en esta comunicación están enmarcados en un proyecto más amplio, que es la evaluación del programa de prácticas en empresa, gestionado por ADEIT durante el curso académico 1991/92. Esta evaluación se realiza todos los años a través de una metodología de encuesta y sus resultados se hacen públicos en un acto oficial organizado por ADEIT en el que intervienen diferentes representantes de los estudiantes, la comunidad universitaria, empresarial y autoridades locales. En este trabajo se presenta un estudio más detallado de los datos más significativos del impacto del programa de prácticas de la asignatura «*Diagnóstico y Tratamiento de las Dificultades de Aprendizaje*», del practicum del Dpto. MIDE. Por ello, en algunos casos, se realizarán comparaciones con el resto de prácticas de la universidad (farmacia, económicas, etc.). Concretamente, los resultados que se analizan son las valoraciones que efectúan tanto los propios estudiantes como los colegios y gabinetes de orientación en donde se realizan las prácticas, en términos de niveles de aprovechamiento, formación, satisfacción, inserción laboral y de orientación profesional. Estas opiniones lógicamente redundarán en una mejora estructural y organizativa de este prácticum.

INSTRUMENTOS

ADEIT ha confeccionado dos cuestionarios para recoger las opiniones tanto de los Estudiantes como de las Empresas. El **primero** está compuesto por 32 ítems agrupados en cuatro grandes bloques: **I)** Datos de identificación del estudiante; **II)** identificación de la empresa receptora; **III)** características de la práctica; compuesto a su vez por los apartados A) características generales de la práctica; B) programa de formación; y C) valoración de la práctica y, finalmente, **IV)** inserción profesional. El segundo cuestionario para las **empresas** incluye 28 cuestiones, además de 4 correspondientes a datos técnicos de la encuesta, estructuradas en 4 apartados: **I)** valoración de la experiencia; **II)** relaciones Universidad-Empresa; **III)** inserción profesional; y **IV)** perfil del futuro estudiante de prácticas.

MUESTRA

Los **estudiantes** entrevistados fueron 647 de un total de 1.162 de la Universitat de València que realizaron prácticas externas en empresas y entidades. El nivel de significación es altamente satisfactorio ya que supera el 0.001. Más significativo resulta este índice en nuestro programa si se considera que de un total de 86 estudiantes del Depto. MIDE contestaron 69. Las **empresas** entrevistadas fueron 176 de un total de 60, lo que indica también su alta representatividad. Lo mismo ocurre con los centros docentes de nuestro programa que contestaron 33 de un total de 43.

APLICACIÓN

Para garantizar la fiabilidad de las entrevistas ADEIT optó por dos estrategias: a) realizar las entrevistas una vez pasado el curso académico 1991/92, entre los meses de noviembre de 1992 hasta enero de 1993 y, b) encargar esta tarea a 8 técnicos especializados, que la realizaron de una forma intensiva y constante. Las encuestas a estudiantes fueron aplicadas indistintamente mediante entrevistas telefónicas desde el ADEIT y entrevistas personales en su sede social. En el caso de las empresas, bastantes de ellas se realizaron en las propias empresas tanto al personal directivo como a los responsables de la formación.

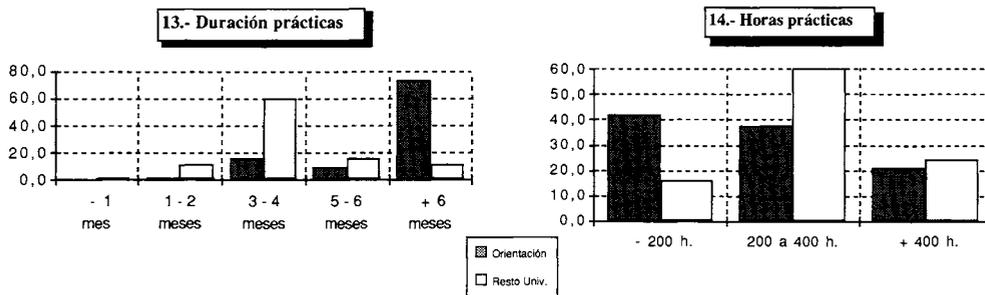
ANÁLISIS DE DATOS

Se ha realizado un análisis descriptivo de los ítems de los dos cuestionarios, especificando en cada uno la distribución de frecuencias total y la correspondiente a cada subgrupo que compone la muestra. Esto es, en el cuestionario de **estudiantes** se pueden distinguir 16 categorías que corresponden con distintas licenciaturas y diplomaturas y en el cuestionario de **empresas** se han diferenciado las siguientes tipologías: Empresas, Farmacias, Ayuntamientos, Centros de Formación y Hospitales. Para este trabajo, se ha realizado una tabulación específica para los estudiantes de nuestro programa, que normalmente realizan el prácticum en *Centros de Formación y Ayuntamientos*. Asimismo, en ambos cuestionarios se han realizado tabulaciones cruzadas entre diferentes ítems al objeto de poder explorar la existencia de relaciones entre variables.

RESULTADOS (I): ENCUESTA ESTUDIANTES

Características generales de la práctica

Previamente, conviene describir algunas cuestiones referidas a la duración de las prácticas (meses, horas, dedicación semanal) ya que, aunque sean notas en principio muy homogéneas para nuestros estudiantes, podemos compararnos con las de otros estudiantes de la universidad y así diferenciar, en relación a otras cuestiones, estas características de las prácticas con la valoración global emitida. Un rasgo característico de nuestras prácticas es que se realizan a lo largo de todo el curso académico (durante más de 6 meses, el 73.5%) con un tiempo estimado que se sitúa alrededor de las 200 horas por curso (200 horas: el 41.8% y entre 200 y 400: el 37.3%). El resto de estudiantes de la universidad realiza fundamentalmente de una forma intensiva, entre 3 y 4 meses (59.7%) y con un tiempo global mayor (entre 200 y 400 horas: el 59.9% y más de 400 horas: el 24%).

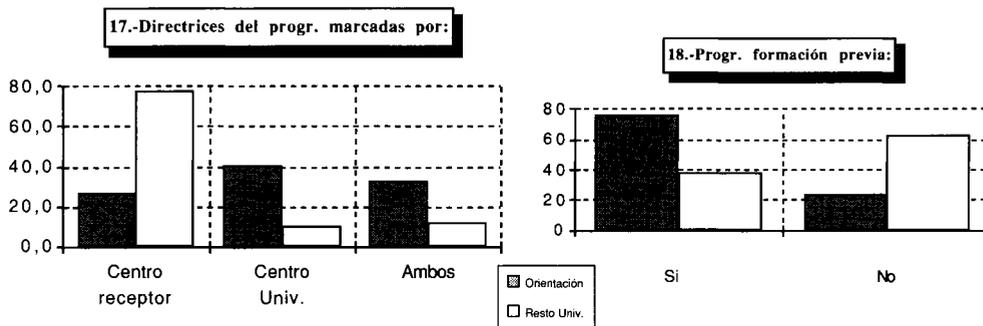


De los cruzamientos efectuados de estos ítems con la valoración global de las prácticas que realizan los estudiantes no se aprecian diferencias significativas. Teniendo en cuenta que el tiempo dedicado por nuestros estudiantes es considerablemente inferior al resto, consideramos que es importante que las prácticas de nuestros alumnos se distribuyan a lo largo de todo el curso con el fin de que puedan valorar la totalidad del proceso de «orientación» que se realiza en los centros educativos, ya que estas variables (duración y horas) no inciden de forma directa en la eficacia del programa.

PROGRAMA DE FORMACIÓN

Este bloque de ítems se relaciona con aspectos estructurales de elaboración del programa de formación desarrollado en las prácticas. En este apartado se aprecian las diferencias más acusadas entre nuestro programa y el del resto de estudiantes. Así, en respuesta al ítem 17, referido a quien marca principalmente las directrices del programa, nuestros estudiantes entienden en un 40.3% que es desde el centro universitario y en un 32.8% que se realiza de una forma coordinada entre el centro universitario y el centro receptor. Por el contrario, la mayoría del resto de estudiantes (77.9%) entienden que las directrices del programa son marcadas por el centro receptor. (*Gráfica ítem 17*).

Esa impresión de nuestros estudiantes, ciertamente ajustada a la realidad, que atribuye al centro universitario como principal diseñador del programa, viene acompañada de otra consideración efectuada sobre la «existencia de un programa de formación previamente definido»: el 76.1% opina que existe, frente a sólo el 37.5% que opina lo mismo en el resto de estudiantes (*Gráfica ítem 18*).

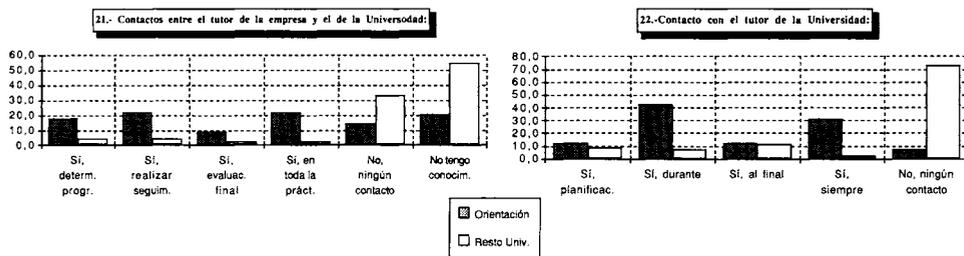


Entre los sujetos que efectuaron sus prácticas a partir de un programa prefijado se produce en todos los casos una valoración positiva sobre el nivel de cumplimiento de las actividades allí incluidas, volviéndose a producir una valoración superior de los estudiantes de nuestro programa (92.0% de los casos) en relación al resto (85.4%).

Resulta también muy significativo la percepción que tiene el alumno sobre el nivel de contactos existente entre el tutor de la empresa y el de la Universidad a lo largo de la práctica. Así, mientras que en nuestro el 70.6% reconoce algún tipo de contacto, en el resto de estudiantes sólo un 13.7% opinan lo mismo (*Gráfica ítem 21*).

También debemos señalar que nuestro programa de prácticas se caracteriza por la dedicación que el tutor de Universidad reserva al propio estudiante, de forma que casi todos los estudiantes reconocen dicha labor, y sólo en un 7.4% de los casos entienden que no han contactado con éste en materia relacionada con las prácticas. Sin embargo, y como contraste con el total, esta cifra se eleva a un 72.8% en el resto de programas. De donde se desprende la implicación tan importante que asume el tutor de la Universidad en el desarrollo de las prácticas profesionales en los centros de trabajo (*Gráfica ítem 22*).

Para finalizar este apartado conviene resaltar que las prácticas realizadas tienen una clara inciden-



cia en la calificación de ciertas asignaturas (69.6%), siendo ello mucho más esporádico en los otros estudios (10.6%). Este fenómeno parece justificar la mayor implicación que asume el profesor en nuestro programa tanto en el diseño como en el desarrollo de las prácticas profesionales.

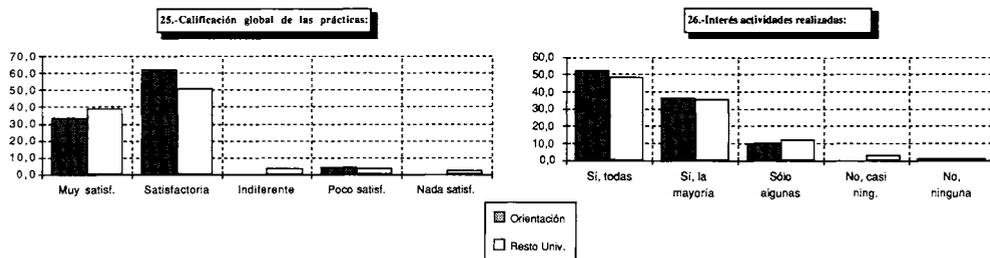
Estos resultados tan satisfactorios pueden deberse a dos factores: a) La larga experiencia del Departamento MIDE en la realización de las prácticas, que se remonta al curso académico 1972-73, y b) La inclusión de una calificación de las prácticas en la evaluación final de ciertas asignaturas.

Valoración de las Prácticas

Globalmente, nuestros alumnos manifiestan una valoración de las prácticas altamente positiva (95.7%), aunque con diverso grado de satisfacción. Esa calificación no difiere demasiado de la opinión del resto de alumnos de otros estudios (90.5%), aunque sea algo superior (media de 4.25 frente a 4.22) (Gráfica ítem 25).

Esa ligera predominancia se mantiene en el ítem relacionado con el interés por las actividades realizadas. En nuestro programa hay un 88.4% de estudiantes que piensan que han sido interesantes todas (52.2%) o la mayoría (36.2%), mientras que en el resto de programas se alcanza un 84.1% (48.6% en todas y 35.5% en la mayoría) (Gráfica ítem 26).

Cuestionados los estudiantes sobre cómo colabora la práctica para la mejora de su formación (se



aceptan hasta dos elecciones), en nuestro programa apuntan como principal alternativa la mayor información profesional que adquieren (66.7%, y 71.8% en el resto de estudiantes). En un segundo nivel, las prácticas contribuyen a clarificar cuestiones teóricas (36.2%, y 22.7% en el resto), así como a dar sentido y coherencia a los estudios realizados (33.3%, y 27.0% de los demás). Finalmente, ninguno de nuestros estudiantes considera que las prácticas no le hayan ayudado a mejorar en algún aspecto concreto (frente al 5.7% del resto de estudiantes). Apreciamos, pues, en la totalidad de estudiantes una cierta compensación que se distribuye entre las diferentes alternativas de mejora de su conocimiento profesional.

Finalmente se ha pedido a los estudiantes que destacaran o clasificaran las actividades desarrolla-

das durante las prácticas. En nuestro caso hay una distribución casi homogénea entre conocimiento del entorno laboral, 33.8% (47.6% en el resto); formación profesional básica, 32.4% (26.5% en el resto); y perfeccionamiento profesional, 32.4% (23.3% en el resto). Otras opciones pueden despreciarse prácticamente.

Inserción Profesional

Si hasta ahora hemos visto aspectos positivos y satisfactorios del programa desarrollado, llegamos a un punto en el que se detecta cierto déficit respecto a otros programas. Ante la pregunta sobre ofertas de continuar en la empresa, sólo a un 2.9% se le ha ofrecido un contrato profesional —laboral o mercantil— y a un 31.9% alguna otra alternativa como es, realizar otras prácticas. Para el resto de programas la primera cifra se eleva hasta un 9.1% de los casos.

A pesar de ello, los estudiantes de nuestro programa consideran que las prácticas son útiles, al menos, para encontrar trabajo relacionado directamente con las actividades realizadas (4.4%); para ayudarles a centrar sus expectativas de empleo en el campo profesional (55.9%, frente a un 40.5% del resto) o para orientar sus expectativas de empleo hacia otros campos profesionales (19.1%). Probablemente la ligera diferencia en la segunda alternativa entre nuestros estudiantes y el resto tenga alguna explicación relacionada con el rol profesional desempeñado y, consecuentemente, con la falta de especificidad de los actuales planes de estudio.

Por otra parte, a esta consideración de las prácticas como elemento fundamental para centrar sus propias expectativas de empleo, se le puede unir los datos de otro ítem que reflejan que esta experiencia les ha servido para ampliar y abrir nuevas perspectivas de empleo (68.2%).

RESULTADOS (II): ENCUESTA EMPRESAS

Nos interesa comentar sólo los resultados de algunas cuestiones recogidas en el apartado de valoración de la experiencia. En general, parece que para los *Centros de Formación*¹ resulta todavía más positiva la experiencia del practicum que para los mismos estudiantes. Señalar que las diferencias de opinión de nuestras «empresas» con el conjunto total de empresas es muy escaso y poco significativo².

Casi unánimemente (94%) los orientadores o directores de los centros piensan que los estudiantes en prácticas han adquirido nuevos conocimientos y todos entienden que el programa de «Prácticas en Empresa», desde el punto de vista de la formación universitaria, es muy necesario (66%) o necesario (33%). Asienten sobre el cumplimiento del programa de formación previsto al inicio de la práctica (94%), siendo en este punto más unánimes que el resto de empresas (84%), confirmándose la tendencia que ya se mostraba claramente en la encuesta de estudiantes respecto a la programación del practicum.

También piensan, la inmensa mayoría, que para sus empresas la presencia del estudiante ha sido satisfactoria (94%), de forma tal como ya se esperaba (73%) o por encima de las expectativas (21%). Entienden, además, que la admisión de los alumnos ha supuesto un esfuerzo en tiempo y profesionales que se ha visto recompensado (88%). Por otra parte, en las prácticas no han tenido ningún problema con los estudiantes (64%), aunque si los ha habido, ha sido por motivo del nivel de formación de los estudiantes (21%) o por su dificultad para adaptarse a la empresa (9%). Es importante resaltar que la mayoría de empresas que colaboran con los diferentes programas de prácticas de la Universitat de València encuentran estas de gran utilidad, condición necesaria para el buen desarrollo de los mismos. En

1 Recuérdese que nuestras «empresas» han sido agrupadas bajo el epígrafe *Centros de Formación*.

2 Se encuentran algunas diferencias significativas entre los distintos grupos de empresas en virtud de su naturaleza y la de los estudiantes que reciben (farmacias, ayuntamientos, hospitales,...), pero ello no es objeto de estudio aquí.

nuestro caso, los motivos que les han inducido a participar en el programa de prácticas (podían elegir más de una opción) parecen ser muy altruistas: Un 79% para proporcionar un complemento a los conocimientos teóricos de los estudiantes; un 30% para facilitar su inserción laboral; un 27% por mantener un contacto con la Universidad y un 6% por la necesidad de contar con personal universitario que colabore con su empresa. A nadie le indujo la motivación de detectar futuros profesionales para la empresa.

CONCLUSIONES

A pesar de no exigir demasiadas horas de permanencia en los centros de trabajo, el hecho de extender las prácticas a lo largo de todo el año académico aventura la hipótesis de que esta dilación es un elemento importante en el aprovechamiento de las mismas. El estudiante va aplicando paulatinamente sus conocimientos teóricos a la práctica, a la vez que integra sosegadamente ésta dentro de su programa de formación universitaria.

Es muy alto el nivel de implicación del profesor-tutor de Universidad en el programa de prácticas, frente al característico de otras experiencias. Este nivel de implicación se traduce en el diseño del programa y en el contacto permanente realizado con los tutores de la empresa y los propios estudiantes durante el desarrollo de las prácticas.

La valoración que efectúan los estudiantes del programa de prácticas es altamente positiva y, si bien en nuestro caso hay una tendencia ligeramente superior, se trata sin duda alguna de una nota característica de las distintas experiencias de prácticas en toda la universidad. Por otra parte, frente a otros programas que sirven para proporcionar mayor información profesional y conocimiento del entorno laboral, nuestros estudiantes aprovechan también las prácticas para lograr una mayor integración y coherencia con los estudios realizados, clarificar cuestiones teóricas y lograr, en definitiva, una mejor formación profesional, tanto básica como aplicada. Finalmente, si bien el programa no contribuye a proporcionar empleo inmediato en el lugar de prácticas, sí que sirve para centrar las expectativas de empleo en el mismo ámbito profesional, a la vez que abre nuevas perspectivas de empleo.

Por lo que respecta a los orientadores y directores de los centros en donde se han efectuado las prácticas, su valoración es también muy satisfactoria, tanto en relación a sus propios centros, como por la ayuda que ofrecen a los estudiantes.

En definitiva, la buena acogida de nuestro programa, tanto por parte de los estudiantes como por los centros, responde a la experiencia dilatada de nuestro programa y a la consideración de la calificación final.

BIBLIOGRAFÍA

- GOODMAN, J. (1993): «La formación de orientadores en Estados Unidos.» En *El compromiso de la Orientación Escolar y Profesional con los cambios de la sociedad*. V Seminario Iberoamericano de Orientación. AEOEP, Madrid, pp. 174-177.
- ECHEVERRÍA SAMANES, B. (1993): «¿Nuevas cualificaciones del orientador o nueva forma de interpretar la profesión.» En *El compromiso de la Orientación Escolar y Profesional con los cambios de la sociedad*. V Seminario Iberoamericano de Orientación. AEOEP, Madrid, pp. 177-188.
- MIDE (1992): *Programa de Prácticas*. Universitat de València: Departamento de Métodos de Investigación y Diagnóstico en Educación.
- REPETTO TALAVERA, E. (1993): «Perspectivas de la formación de orientadores. El problema de la capacitación en una dimensión europea.» En *El compromiso de la Orientación Escolar y Profesional con los cambios de la sociedad*. V Seminario Iberoamericano de Orientación. AEOEP, Madrid, pp. 30-39.

EVALUACIÓN DE LOS EFECTOS DEL PROGRAMA DE ORIENTACIÓN VOCACIONAL Y PROFESIONAL «MIRANDO HACIA EL FUTURO»

M^a del Mar Elías Caraballo y M^a de los Ángeles Lorenzo Miralles
Universidad de Sevilla

INTRODUCCIÓN

La evaluación de nuestro Programa de Orientación Vocacional y Profesional «Mirando hacia el futuro» se sitúa en una perspectiva técnica. Con ella se pretende obtener información documentada sobre los efectos del programa con el fin de demostrar la rentabilidad del programa y comprobar si éste da respuestas y satisfacción a las necesidades sentidas por los alumnos.

La evaluación va dirigida al conjunto de profesores del C.P. Padre Manjón de Bormujos (Sevilla), cara a una **toma de decisiones** fundada a partir de la valoración del proceso de implementación del programa y a partir de los logros obtenidos en el mismo.

En coherencia a las cuestiones señaladas anteriormente, hemos seleccionado el modelo CIPP de Stufflebeam, asumiendo que la evaluación es «el proceso de identificar, obtener y proporcionar información útil y descriptiva acerca del valor y el mérito de las metas, la planificación, la realización y el impacto de un objeto determinado, con el fin de servir de guía para la toma de decisiones, solucionar los problemas de responsabilidad y promover la comprensión de los fenómenos implicados» (Stufflebeam, 1989, pp. 183).

Si bien se ha realizado la evaluación del proceso y del producto, esta exposición se va a centrar exclusivamente en los efectos del programa.

PRINCIPALES CARACTERÍSTICAS DEL PROGRAMA DE ORIENTACIÓN EVALUADO

Las notas fundamentales que caracterizan el programa son las siguientes:

— Su finalidad general es el desarrollo de la madurez vocacional. Especialmente se pretende la mejora en habilidades de toma de decisiones, autoconocimiento, incremento de la conducta de búsqueda de información y reducción de los estereotipos sobre las distintas carreras y profesiones.

— Suministra información a dos niveles: información referida al propio sujeto (intereses, actitudes, valores, etc.) e información ocupacional y profesional. Este último nivel proporciona información acerca de aquellas profesiones que mayor interés han mostrado de las existentes en el mercado laboral; sobre las instituciones educativas; e identificación de fuentes generales de información.

En definitiva el programa responde a un enfoque comunitario, centrado en el sujeto y arraigado en

el medio. Es predominantemente activo y multidimensional, atendiendo especialmente a las características evolutivas de los sujetos y a sus etapas de desarrollo. Desde éste, la orientación se concibe como ayuda al proceso de desarrollo del individuo en el ámbito de sus experiencias educativas y vocacionales, centrando su atención en la función del yo y en el concepto de sí mismo. Se estructura como un proceso de ayuda al estudiante en su desarrollo hasta alcanzar su madurez. Entendiéndose el desarrollo de éste como efecto de una interacción activa de él mismo con el medio.

PROBLEMA DE INVESTIGACIÓN

El problema planteado en la investigación es el siguiente:

Conocer los efectos causados por el Programa de Orientación «Mirando hacia el futuro» en el desarrollo de la madurez vocacional en los alumnos de 8º curso de EGB del C.P. Padre Manjón de Bormujos (Sevilla).

HIPÓTESIS DE TRABAJO

— Hipótesis Principal:

La implementación del Programa implica un aumento significativo del nivel de madurez vocacional y profesional del alumno.

— Subhipótesis:

1.— Tras la implementación del Programa existirá una variación en las actitudes ante la elección profesional, en los siguientes sentidos: Los alumnos participarán más activamente en el proceso de hacer elecciones; Los alumnos no se apoyarán en otros en la elección de una ocupación; Los alumnos basarán su elección en varios factores; Los alumnos poseerán concepciones precisas acerca del proceso de elección ocupacional.

2.— Tras la implementación del Programa los alumnos poseerán un conocimiento mayor del número de las ocupaciones existentes en diferentes áreas ocupacionales.

3.— Tras la implementación del Programa los alumnos poseerán un mayor conocimiento de las características relevantes de las distintas profesiones.

4.— Los alumnos, tras la implementación del Programa, identificarán los pasos requeridos para ingresar en cualquier profesión.

5.— Los alumnos, tras la implementación del Programa elegirán un trabajo teniendo en cuenta las características de sí mismo.

DISEÑO

Se ha seleccionado un «Diseño cuasiexperimental de muestras separadas pretest-postest».

TABLA 1

R	O ₁	(X)	
R		X	O ₂

Aunque inicialmente se pretendía analizar sólo los efectos producidos por el programa en el curso de 8º de forma global, una primera aproximación a la realidad de las aulas puso de manifiesto la

conveniencia de estudiar los resultados en cada grupo clase. De esta manera, se controlaba la variable implementador.

MUESTRA

La población está compuesta por 90 sujetos cuyas edades oscilan entre los 13 y los 15 años. Están divididos en tres grupos-clase de caracteres homogéneos.

De acuerdo con el diseño de investigación hemos seleccionado dos muestras diferentes a través de un procedimiento probabilístico **aleatorio sistemático**.

PROCEDIMIENTO E INSTRUMENTOS DE RECOGIDA DE DATOS

Los datos se han recogido a través de los siguientes instrumentos:

— **Cuestionario de intereses profesionales, CIPSA** para medir la variable «Información ocupacional cuantitativa». Con el fin de identificar el número de profesiones que conoce el individuo en las distintas áreas ocupacionales.

— **Cuestionario de Madurez Vocacional** es una adaptación del CMI de Crites en fase de validación por Espín, J. y otros de la Universidad de Barcelona. Se ha diseñado con el fin de determinar los siguientes aspectos relacionados con la elección vocacional y profesional: Actitud de los alumnos hacia la elección vocacional; Grado en que un individuo conoce las características relevantes de las distintas ocupaciones; Capacidad de planificación; Capacidad del sujeto para seleccionar la profesión más adecuada, en función de las distintas ocupaciones.

Esta prueba consta de dos bloques:

Una escala de actitudes, «Actitudes ante la elección de un trabajo o profesión» mediante la cual mediremos las variables siguientes: Implicación en el proceso de elección, Independencia en la toma de decisiones, Preferencia por factores de elección vocacional y Concepciones del proceso de elección.

Un cuestionario sobre una serie de competencias necesarias para la elección de un trabajo o profesión. Este bloque a su vez consta de las siguientes partes: Conociendo trabajos, Escogiendo un trabajo y Mirando hacia adelante. A través de dicho cuestionario mediremos las variables: Información ocupacional cualitativa, Selección de metas y Planificación.

ANÁLISIS DE LOS DATOS

Se ha realizado un análisis cuantitativo de los datos a través del paquete estadístico SPSS/PC+. Para contrastar las dos mediciones realizadas (pretest-postest) se han utilizado las siguientes pruebas:

— La prueba *t* de Student cuando estudiamos las diferencias significativas globales (sin diferenciar por cursos) entre las dos mediciones.

— La *U* de Mann-Whitney cuando estudiamos las diferencias entre las dos mediciones por curso.

Los datos obtenidos de la **T de Student** los recogemos en la tabla 2.

Consideramos como hipótesis nula (H_0) la no existencia de diferencias significativas entre los resultados obtenidos entre las dos medidas (pretest y postest) y que el grado de significación, α , es de 0,05. Los resultados demuestran que las variables en las que ha incidido el programa han sido el conocimiento de profesiones y la capacidad de planificación.

TABLA 2

	T	p
V1 IMPLICACIÓN	-0.72	0.472
V2 INDEPENDENCIA	0.47	0.641
V3 FACTORES	0.91	0.367
V4 PRECISIÓN	1.38	0.171
V5 PROFESIONES	-3.52	0.001
V6 CONOCER	-0.62	0.538
V7 PLANIFICAR	-5.91	0.000
V8 ESCOGER	-1.57	0.122

Los datos que hemos obtenido de la **U de Mann-Whitney** los recogemos en las siguientes tablas:

TABLA 3

CURSO 1			
VARIABLES	U	Z	CORRECTED P
V1 IMPLICACIÓN	74.5	-0.1408	0.8881
V2 INDEPENDENCIA	70.0	-0.3904	0.6962
V3 FACTORES	68.5	-0.4716	0.6372
V4 PRECISIÓN	62.5	-0.8044	0.4212
V5 PROFESIONES	43.0	-1.8631	0.0624
V6 CONOCER	75.0	-0.1147	0.9087
V7 PLANIFICAR	37.5	-2.2092	0.0272
V8 ESCOGER	54.5	-1.2640	0.2062

Si observamos la tabla 3 comprobamos que la variable 7 (Planificar) en el **curso 1**, con un valor de $U= 37.5$, es en la única en la que se puede rechazar la hipótesis nula siendo $P= 0.0272$. También en el curso 2 (tabla 4) la variable 7 (Planificar) es la única en la que se puede rechazar la hipótesis nula, con un valor de $U= 27.0$, siendo $P= 0.0133$. En el curso 3 (tabla 5) son las variables 5 (Profesiones) y 7 (Planificar) en las que se puede rechazar la hipótesis nula, con unos valores de U de 17.0 y 10.5, respectivamente, siendo $P=0.0045$, en el primer caso y $P= 0.0008$ en el segundo. En los tres casos existen diferencias significativas entre el pretest y el postest en estas dos variables.

TABLA 4

CURSO 2			
VARIABLES	U	Z	CORRECTED P
V1 IMPLICACIÓN	50.5	-0.9843	0.3249
V2 INDEPENDENCIA	45.5	-1.3036	0.1924
V3 FACTORES	57.5	-0.5361	0.5919
V4 PRECISIÓN	43.5	-1.4218	0.1551
V5 PROFESIONES	43.0	-1.4173	0.1564
V6 CONOCER	58.0	-5.239	0.6004
V7 PLANIFICAR	27.0	-2.4754	0.0133
V8 ESCOGER	61.0	-0.3134	0.7540

TABLA 5

CURSO 3			
VARIABLES	U	Z	CORRECTED P
V1 IMPLICACIÓN	45.5	-1.0025	0.3161
V2 INDEPENDENCIA	55.0	-0.3363	0.7359
V3 FACTORES	48.5	-0.7703	0.4411
V4 PRECISIÓN	58.5	-0.1003	0.9201
V5 PROFESIONES	17.0	-2.8377	0.0045
V6 CONOCER	53.0	-0.4739	0.6356
V7 PLANIFICAR	10.5	-3.3693	0.0008
V9 ESCOGER	44.5	-1.0515	0.2930

CONCLUSIONES

El análisis de los datos permite confirmar que los alumnos identifican mejor la secuencia correcta de pasos requeridos para ingresar en una determinada ocupación y conocen un número mayor de ocupaciones.

Aunque estos resultados pueden considerarse insuficientes, nosotros nos sentimos satisfechos debido a que los cambios más significativos se han dado en la Planificación y Conocimiento de las profesiones, dimensiones identificadas como más relevantes en el contenido del programa. Por tanto, el progreso en la capacidad de planificación se estima como muy positivo, ya que el programa pretendía no sólo que los alumnos se decidían por unos estudios o una profesión determinada, sino también que sepan tomar decisiones razonadas y coherentes. Estas necesitan ser planificadas con anterioridad para que la decisión final posea una estructura lógica que posibilite una puesta en práctica con visos de éxito.

Como hemos dicho al comienzo, también ha aumentado el número de ocupaciones que en un principio conocían los alumnos. Esto supone un logro importante del programa ya que las expectativas

profesionales iniciales de los alumnos eran muy restringidas centradas casi exclusivamente en el nivel de estudios de Formación Profesional.

BIBLIOGRAFÍA

- ÁLVAREZ ROJO, V. (1987): *Metodología de la orientación educativa*. Alfar, Sevilla.
- ALVIRA, F. (1991): *Metodología de la evaluación de programas*. Centro de Investigaciones Sociológicas. Cuadernos metodológicos, n. 2. Madrid.
- BERK, R. y ROSSI, P. H. (1990): *Thinking about Program Evaluation*. Sage Publ., Londres.
- BISQUERRA, R. (1987): *Introducción a la estadística aplicada a la investigación educativa*. P.P.U., Barcelona.
- CAMPBELL y STANLEY (1988): *Diseños experimental y cuasiexperimental en la investigación social*. Amorrortu, Buenos Aires.
- COLÁS, M. P. y BUENDÍA, L. (1992): *Investigación educativa*. Alfar, Sevilla.
- COLÁS, M. P. y REBOLLO, M. A. (1993): *Evaluación de Programas. Una guía práctica*. Ed. Kronos, Sevilla.
- ETXEBERRIA, J. (1990): *Programación y análisis estadístico básicos con spss-pc+*. Paraninfo, Madrid.
- FITZ-GIBBON, C. T. & MORRIS, L. L. (1987): *How to design a program evaluation*. Sage, Beverly Hills, CA.
- MADAUS, G. F., SCRIVEN, M. S., y STUFFLEBEAM, D. S. (1988): *Evaluation models: viewpoints educational and human services evaluation*. Kluwer-Nijhoff, Boston.
- RIVAS (1988): *Psicología vocacional. Enfoques del asesoramiento*. Morata, Madrid.
- SANZ ORO, R. (1990): *Evaluación de programas de orientación educativa*. Pirámide, Madrid.
- RODRÍGUEZ, M. L. y GRAS, M. (1986): *Modelos de orientación profesional en el aula*. Oikos-tau, Barcelona.
- SALVADOR, A. y PEIRO, J. M. (1986): *La madurez vocacional*. Ed. Alhambra, Madrid.
- STUFFLEBEAM, D. y SHINKFIELD, A. (1987): *Evaluación sistemática*. Ed. Paidós, Barcelona.
- WORTHEN, B. R. & SANDERS, J. R. (1987): *Educational Evaluation. Alternative approaches and practical guidelines*. Longman Inc. London.

COMUNICACIONES PONENCIA III

ERGO: CARACTERÍSTICAS Y APLICACIONES

I. Inmaculada Asensio Muñoz
Universidad Complutense de Madrid

En el contexto de las inferencias diagnósticas que se definen en la ponencia presentada por Orden, Gaviria, Fuentes y Lázaro (1993) cobra interés el estudio del «software» desarrollado para facilitar este trabajo. Este es el marco en el que se sitúa nuestra comunicación, puesto que ERGO es un programa informático desarrollado por Noetic Systems con el objetivo de servir de herramienta en procesos de definición de redes de probabilidad y de estimación de probabilidades de estados latentes. Es un soporte técnico para la toma de decisiones aplicable a distintos campos, entre ellos el educativo.

CARACTERIZACIÓN

ERGO es un entorno de trabajo gráfico para la construcción de sistemas expertos de probabilidad con Macintosh. Es un programa que facilita la creación, desarrollo, depuración y evaluación de redes probabilísticas. Estas redes se representan a través de gráficos no recursivos en los que los nodos son las variables y los vectores significan asociación o relación entre dichas variables.

Constituyen el programa tres elementos básicos: un entorno gráfico de dibujo —**editor de redes**—, al que acompaña un editor para cuantificar la relación entre las variables —**editor de probabilidades**— y un **motor de inferencias lógico-probabilísticas** para calcular las probabilidades derivadas del modelo.

Con estas herramientas, el usuario simplemente ha de definir las variables de interés y la importancia de las asociaciones entre las variables a través del grafo, que puede ser realizado directamente en el ordenador. Esta tarea se simplifica casi hasta convertirse en una actividad lúdica cuando se realiza en un entorno Macintosh. El resultado es una red de relaciones entre variables perfectamente clara e inteligible a primera vista.

Por otra parte se han de especificar las probabilidades asociadas con cada variable definida, así como las probabilidades condicionadas. El editor de probabilidades de ERGO proporciona medios para probar la consistencia de las mismas y para normalizar distribuciones.

ERGO utiliza algoritmos basados en la teoría de grafos y en la teoría de probabilidad para chequear y confirmar la consistencia interna de las redes elaboradas, tanto durante la construcción del sistema como durante la inferencia (Lauritzen y Spiegelhalter, 1988).

FUNCIONAMIENTO

Para ejemplificar el proceso que se sigue en la utilización de este programa vamos a utilizar un supuesto sencillo (NOETIC SYSTEMS INCORPORATED, 1992).

Partimos de que en un estudio diagnóstico interesan fundamentalmente tres variables: «tuberculosis», «cáncer de pulmón» y «bronquitis». Son dos hechos conocidos a) que fumar se asocia con las dos últimas enfermedades y b) que haber visitado recientemente Asia se relaciona con el padecimiento de la tuberculosis de manera que en personas que han realizado este viaje se incrementa la probabilidad de contraer dicha enfermedad. Toda esta información se transmite al programa de modo gráfico, esto es, dibujando cinco círculos o nodos (ASIA, TABACO, TUBERCULOSIS, CÁNCER, BRONQUITIS) y tres vectores relacionales (uno que va de ASIA a TUBERCULOSIS y dos que salen de TABACO y van a CÁNCER y a BRONQUITIS).

Además el experto sabe que tanto la tuberculosis como el cáncer de pulmón (TUBoCA) dan lugar a radiografías donde se manifiesta alguna anomalía y las tres enfermedades descritas se manifiestan con disnea o dificultad respiratoria. Esta nueva información se incluye gráficamente añadiendo tres nodos (TUBoCA, RADIO, DISNEA) y cinco vectores (de TUBERCULOSIS y de CÁNCER a TUBoCA, de TUBoCA a RADIO y DISNEA y de BRONQUITIS a DISNEA). Con esto se termina de dibujar la red de relaciones que aparece en la pantalla del ordenador.

Abriendo la ventana de probabilidades se procede posteriormente a la definición de cada variable. Para ello se escribe su denominación, se indica el número de valores que puede tomar y se da nombre a cada uno de dichos valores (si la variable es dicotómica se pueden definir como PRESENCIA/AUSENCIA, ACIERTO/ERROR, NORMAL/ANORMAL...). También aquí se pueden modificar las probabilidades condicionales para la variable en cuestión. El editor de probabilidades permite copiar y pegar datos procedentes de hojas de cálculo, otros nodos o de un fichero de datos. El programa comprueba automáticamente la validez de la matriz de probabilidades independientemente de cuál sea la fuente de la que proceden. Por último ERGO compila las redes para calcular las probabilidades principales en todos y cada uno de los nodos definidos, que se presentan gráficamente como histogramas. A partir de este momento se puede proceder al cálculo de probabilidades en variables asociadas para un caso concreto. Por ejemplo, la probabilidad de que padezca cáncer de pulmón una persona que ha dado una radiografía anormal o de la que se conocen o no otros datos acerca de otras variables relevantes.

Todo este proceso así explicado parece muy laborioso. Sin embargo, con una demostración en la que se disponga de un ordenador Macintosh se puede entender fácilmente la secuencia y la simplicidad que supone el servirse de este instrumento.

Un ejemplo con contenido educativo es el que se describe en la ponencia presentada por Oorden, Gaviria, Fuentes y Lázaro (1993). En él se explica cómo, definidas seis situaciones de prueba o ítems diferentes, con una teoría de fondo, pueden interpretarse como «síntomas» —siguiendo la metáfora médica— de un estado latente o variable compleja —constructo— denominado «nivel cognitivo». De manera que dada una red de relaciones entre los ítems y el constructo y conocidas las probabilidades condicionadas de que el estado cognitivo sea «Y» dado que la respuesta a cada ítem ha sido «X1», «X2», «X3», «X4», «X5» y «X6» puede llegar a hacerse el diagnóstico individual de un niño del que se conocen sus respuestas a las seis situaciones de prueba. Puede, por supuesto, utilizarse también el programa para la identificación de tipologías.

APLICACIONES Y VALORACIÓN

De todo lo dicho se deduce que, en nuestro campo como en el campo médico, las aplicaciones

básicas del programa se derivan de las posibilidades que ofrece para el diagnóstico y la evaluación educativas —tanto en la fase de construcción de tests, como en la de interpretación de las puntuaciones obtenidas por los diversos sujetos que responden al test y en la toma de decisiones—.

El uso de este programa en trabajos de medida, diagnóstico, evaluación e investigación educativas puede ser de gran ayuda en lo que supone de simplicidad y economía de tiempo y también es un soporte técnicamente muy útil para la toma de decisiones.

Sin embargo, chocamos con una dificultad de partida que se produce por el hecho de que en nuestro campo no existen teorías fuertes que permitan determinar la red de relaciones entre un conjunto de variables relevantes —es posible que no dispongamos siquiera de una selección de variables relevantes, según el campo concreto de que se trate—. Esta limitación —no de la herramienta informática sino de la disciplina en la que pretende utilizar— supone que los modelos posiblemente hayan de ser inicialmente poco ambiciosos por lo general.

No obstante, este método tiene a nuestro juicio la gran ventaja de que aplicado al proceso de construcción de tests pone al especialista ante el reto de definir estados latentes subyacentes a cada ítem o grupo de ítems. Esto es tanto como decir que favorece la aplicación de principios básicos defendidos por los teóricos de la medida educativa como es la prioridad otorgada a la validez de constructo. Pensamos que este instrumento informático facilita el hecho de que los constructores de tests se sitúen en una teoría o ciencia-marco antes de proceder al trabajo operativo.

BIBLIOGRAFÍA

- LAURITZEN, S. L. y SPIEGELHALTER, D. J. (1988): «Local Computations with Probabilities on Graphical Structures and Their Application to Expert Systems». *Journal of the Royal Statistical Society*, Series B, 50, pp. 157-224.
- MISLEVY, R. J. (1993): «Probability-Based Inference in Cognitive Diagnosis». *Documento presentado en la Conference on Alternative Diagnostic Assessment* Universidad de Iowa, 6-7 mayo.
- NEAPOLITAN, R. E. (1990): *Probabilistic Reasoning in Expert Systems. Theory and Algorithms*. John Wiley & Sons Inc., Nueva York.
- NOETIC SYSTEMS INCORPORATED (1992): *ERGO. Product Information*, Noetic Systems Incorporated, junio.
- ORDEN, A., GAVIRIA, J. L., FUENTES, A. y LÁZARO, A. (1993): *Modelos de Construcción y Validación de Instrumentos Diagnósticos*. Ponencia presentada al VI Seminario de A.I.D.I.P.E., septiembre.
- PEARL, J. (1988): *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann Publ., San Mateo (CA) (2ª Ed.).

APLICACIÓN DE LOS NUEVOS ENFOQUES DE TEORÍA DE LA MEDIDA EN LA CONSTRUCCIÓN DE UNA PRUEBA DIAGNÓSTICA DE COMPRENSIÓN VERBAL

F. Javier Murillo Torrecilla
C.I.D.E. Servicio de Investigación

Una de las grandes contribuciones de los nuevos enfoques psicométricos (Mislevy, 1993; Orden y cols., 1993) es haber acabado con el enfrentamiento tradicional entre las teorías de construcción de test y las últimas propuestas de la psicología cognitiva. Las aportaciones realizadas en ese sentido por Robert J. Mislevy, entre otros, apuestan por una teoría que, superando los enfoques clásicos de la medida y la Teoría de Respuesta al ítem, permita construir pruebas que sean sensibles a las distintas fases de desarrollo y procesos a través de los que el sujeto adquiere una destreza cognitiva. Los fundamentos de estas nuevas ideas pueden encontrarse en los modelos alternativos de medida y los desarrollos en las redes inferenciales basadas en probabilidades. Según Mislevy (1993, 21):

«(...) Meaningful tasks in learning domains generally entail multiple strands of conceptual development, and that a probability-based inference network is an appropriate tool for planning, interpreting, and decision-making in pedagogical context».

De esta manera, los test contruidos con estas ideas permitirán obtener una panorámica de los procesos cognitivos de un sujeto y su nivel de adquisición de las destrezas implicadas en un ámbito cognitivo determinado.

En esta breve comunicación se pretende mostrar una aplicación de los nuevos enfoques psicométricos en la elaboración de una prueba que permita conocer el nivel de adquisición de procesos cognitivos implicados en la comprensión verbal. Para ello, se ha partido de una de las teorías de la comprensión verbal más elaboradas de entre las aparecidas en estos años, teoría de comprensión verbal de R. J. Sternberg y, más concretamente, en la subteoría de la adquisición de componentes. Las propuestas que se realizan constituyen una parte de una investigación más amplia y representan los primeros pasos a dar en la construcción de la prueba.

Sternberg (1987a), partiendo de su conocida teoría triárquica de la inteligencia, elabora una propuesta formada por tres subteorías que abordan otros tantos aspectos del procesamiento de la información verbal: la *subteoría de los componentes de adquisición del conocimiento*, donde explica cómo se desarrolla el aprendizaje de las palabras desconocidas a partir del contexto (Sternberg, 1987b); la *subteoría de los componentes de actuación en comprensión en tiempo real*, donde analiza las destrezas utilizadas en el funcionamiento verbal cotidiano en tiempo real (Sternberg y Mcnamara,

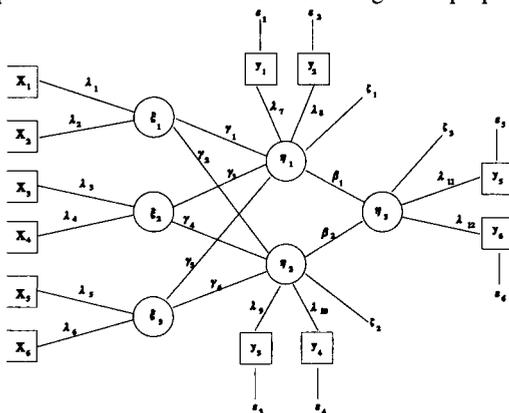
1985); y la *subteoría de los metacomponentes*, donde estudia cómo se utilizan los procesos ejecutivos en la comprensión verbal (Wagner y Sternberg, 1987). Desde un planteamiento puramente educativo interesa fundamentalmente la primera de ellas, dado que, entre otras cuestiones, permite conocer cómo se pueden entrenar destrezas de adquisición de vocabulario.

La subteoría de los componentes de adquisición del conocimiento defiende que son tres los procesos implicados en la extracción del significado de una palabra desconocida a partir del contexto: la codificación selectiva, la comparación selectiva y la combinación selectiva. Estos procesos se aplican a determinados elementos de información, denominados claves contextuales. Claves que están condicionadas por una serie de variables, las llamadas variables mediadoras. Según la teoría, el significado de una palabra desconocida puede ser extraído de los elementos que rodean a la frase (contexto externo) o de los propios componentes de la palabra (contexto interno). Así surge la subteoría de la descontextualización para el contexto externo y para el contexto interno. Para ambas, los procesos son los mismos, sin embargo, las claves y las variables mediadoras varían. Así, la destreza de adquisición de palabras está compuesta de tres procesos o destrezas básicas que el sujeto debe realizar: la codificación selectiva, la combinación selectiva y la comparación selectiva. Cada una de ellas se aplica al contexto interno y al externo. Por lo tanto, un test para diagnosticar el nivel de la destreza de adquisición de componentes debe medir cada uno de esos procesos básicos y su aplicación a un tipo de contexto determinado, a través de la cuantificación de las variables observadas.

Los nuevos enfoques psicométricos proponen la construcción de pruebas diagnósticas a partir de un modelo que recoja los procesos implicados en el dominio de la destreza. Este modelo, que representa la arquitectura cognitiva del sujeto, debe incluir los elementos que forman parte del modelo, sus relaciones estructurales y la operacionalización de esas variables. Los elementos que forman parte del modelo son las tareas (variables observables) y los procesos (variables latentes), así como la operativización de sus relaciones consiste en la determinación de las probabilidades condicionales asociadas. En esos conjuntos de probabilidades se apoya la tecnología de *Bayesian belief networks* (Lauritzen y Spiegharter, 1988). Los sistemas expertos basados en la misma permiten manejar la incertidumbre asociada a los procesos de diagnóstico cognitivo.

De esta manera, parece coherente que el modelo para la elaboración de la prueba pueda corresponderse a un modelo causal basado en la teoría cognitiva de partida, es decir, en este caso la subteoría Sternberg. La determinación de probabilidades asociadas a cada proceso básico que conforma la destreza de la adquisición de componentes de comprensión verbal se realizaría mediante la validación de dicho modelo causal.

El modelo causal para esta subteoría se concreta en la siguiente propuesta:



Donde

- ξ_1 — Codificación selectiva
- ξ_2 — Comparación selectiva
- ξ_3 — Combinación selectiva
- η_1 — Contexto externo
- η_2 — Contexto interno
- η_3 — Comprensión Verbal.

La utilización de este modelo causal como base para la construcción de la prueba diagnóstica de comprensión verbal, plantea una serie de problemas. Uno de ellos es que las variables observables y latentes son continuas, en la forma en que está formulado. Sin embargo, los sistemas expertos desarrollados hasta ahora con esa tecnología de redes de probabilidad únicamente permiten que las variables sean dicotómicas. Este obstáculo puede ser evitado mediante la transformación de estas variables a discretas finitas, a través de una sencilla modificación.

Respecto a los ítems que forman parte del modelo, con las variables mediadoras propuestas por Sternberg, podría elaborarse un sistema de generación de los mismos, de manera análoga a la planteada por Katz (1988).

A través de la mejora de la estructura de los modelos y de la búsqueda de los patrones de generación de ítems se podrían encontrar vías que completen y concreten algunos aspectos de la teoría de la adquisición de componentes de Sternberg. Así, este modelo causal se podría convertir en el núcleo de un sistema experto de diagnóstico de la comprensión verbal, al igual que podría servir para realizar la validación empírica del modelo cognitivo subyacente y la mejora de la teoría que lo genera.

REFERENCIAS BIBLIOGRÁFICAS

- KATZ, B. (1988): *Using English for indexing and retrieving*. Massachusetts Institute of Technology: Cambridge, MA.
- LAURITZEN, S. L. y SPIEGELHALTER, D. J. (1988): Local computations with probabilities on graphical structures and their application to expert systems. *Journal of the Royal Statistical Society*, 50, 157-224.
- MISLEVY, R. (1993): Foundations of a new test theory. En Norman Frederiksen, Robert J. MISLEVY e Isaac BEJAR (Eds.): *Test Theory For a New Generation of Tests*. Lawrence Erlbaum Associates: Hillsdale, N. J. 19-39.
- STERNBERG, R. J. (1987a): The psychology of verbal comprehension. En R. GLASER (Ed.): *Advances in Instrumental Psychology*. Vol. 1. Lawrence Erlbaum Associates: Hillsdale, N. J. 97-151.
- STERNBERG, R. J. (1987b): Most vocabulary is learned from context. En M. MCKEON y M. E. CURTIS (Eds.): *The nature of vocabulary acquisition*. Lawrence Erlbaum Associates: Hillsdale, N. J.
- STERNBERG, R. J. y MCNAMARA, T. P. (1985): The representation and processing of information in real-time verbal comprehension. En S. E. EMBRETSON (Ed.): *Test design. Developments in psychology and psychometrics*. Academic Press: New York, N.Y.
- ORDEN, A. de la, GAVIRIA, J. L., FUENTES, A. y LÁZARO, A. (1993): Modelos de construcción y validación de instrumentos diagnósticos. Ponencia presentada en el VI Seminario de Modelo de Investigación Educativa. Madrid, 23 a 25 de septiembre.
- WAGNER, R. K. y STERNBERG, R. J. (1987): Executive control in reading comprehension. En B. K. BRITTON y S. M. GLYNN (Eds.): *Executive control in processes in reading*. Lawrence Erlbaum Associates: Hillsdale, N. J.

MODELOS LOGIT EN LA VALIDACIÓN Y EL ANÁLISIS DE REGISTROS DE OBSERVACIÓN

Juan Carlos Tójar Hurtado
Facultad de Ciencias de la Educación
Universidad de Málaga

Del proceso de una investigación que utiliza la metodología observacional (v. g. Anguera, 1983, 1991; Bakeman y Gottman, 1986), en este trabajo se van a destacar sólo dos etapas o *momentos* (Tójar, 1990): la *optimización de los registros* y el *análisis de los datos*.

La optimización de los registros está íntimamente relacionada con la calidad de los datos que han sido recogidos tras la aplicación de un instrumento (normalmente un sistema de categorías). Previo al análisis de los datos es preciso determinar si los registros son suficientemente *replicables* (Bakeman y Gottman, 1986). Para contrastar la replicabilidad de los registros se han utilizado conceptos como el acuerdo, la fiabilidad y el resto de características relativas a la medida. En un trabajo reciente (Tójar, 1992), se propone utilizar la *concordancia* de los registros como un concepto capaz de aglutinar diferentes características en la medida observacional. De esta forma, se podría hablar de calidad de la medida observacional cuando un conjunto de registros concuerda (tanto si son realizados por diferentes observadores independientes, como por el mismo en diferentes sesiones, o con diferentes instrumentos...).

Una vez garantizada la calidad de los registros recogidos cobra sentido la etapa del análisis de los mismos. Desde hace un par de décadas, la investigación observacional viene introduciendo la perspectiva secuencial en el registro y también en el análisis de los datos. Esta perspectiva se traduce en la consideración de los parámetros *orden*, e incluso *duración*, además de por supuesto la *frecuencia*, en la descripción de lo observado en un contexto dado. El conjunto de técnicas capaz de abordar datos secuenciales categóricos se conoce con el nombre de *análisis secuencial* de datos observacionales (Bakeman, 1991).

Los objetivos del análisis secuencial se pueden resumir en *descubrir patrones estocásticos en los datos* y *evaluar el efecto que variables contextuales y explicativas en la estructura secuencial* (Gottman y Roy, 1989: 19). Analizando secuencias se pretende, además de la identificación de las situaciones o comportamientos más probables en un contexto dado, la descripción, por un lado, y la explicación y predicción, por otro, de los estados por los que pasa una interacción. La descripción se puede realizar mediante cadenas de secuencias de situaciones previamente categorizadas, y en cada instante es posible predecir qué situaciones son inhibidas y/o excitadas con respecto a una en concreto (criterio).

Las primeras técnicas de análisis secuencial (*cadenas de Markov* o la *técnica de retardos* de Sackett, 1979), han dado paso en los últimos años al uso de enfoques multivariados, principalmente el

análisis log-lineal, que facilitan la construcción de modelos capaces de abarcar las relaciones entre numerosas variables (sistemas de categorías, diferentes codificadores, sesiones, instrumentos, ...).

Existen fundamentalmente dos aproximaciones log(aritmico)-lineales para modelizar las frecuencias en una tabla de contingencia. Por un lado, los *modelos log-lineales*, en los que no existe diferencia entre variables dependientes e independientes (*perspectiva simétrica* en la que todas serían variables respuesta) y, por otro, los *modelos logit*, en los que una variable es considerada como dependiente (*perspectiva asimétrica*).

Como mostraron Bakeman, Adamson y Strisik (1989) la perspectiva asimétrica de los modelos logit es perfectamente compatible con el diseño secuencial. Simplemente basta con situar una variable consecuente (C) en el tiempo $t + k$ como variable dependiente y el resto (v. g. antecedentes en los tiempos $t + k - 1, t + k - 2, t + k - 3, \dots, t$, registros realizados por diferentes codificadores, o en diferentes momentos, o con otros instrumentos...), como independientes.

La perspectiva secuencial utilizada frecuentemente, como se apuntó con anterioridad, en las etapas de registro y de análisis se ha aplicado escasamente a la hora de estudiar la calidad de la medida observacional realizada. Aun cuando se puede reconocer el interés de los trabajos pioneros en este sentido de Gottman (1979, 1980), no se puede considerar una cobertura rigurosa del problema hasta la introducción del concepto de *frecuencia de transición* (Wampold y Holloway, 1983). A partir de una matriz de frecuencias de transición para cada registro, Wampold y Holloway proponen utilizar el *coeficiente de correlación de Pearson* o, en casos más complejos, incluir las frecuencias de transiciones como una faceta de análisis de variancia en el contexto de la *teoría de la generalizabilidad*.

Algunos autores han mostrado la utilidad de las aproximaciones log-lineales para evaluar la concordancia (no secuencial) en matrices de confusión (Tanner y Young, 1985, 1986; Agresti, 1990). Más recientemente en otro lugar (Tójar, 1992) se ha mostrado como los modelos log-lineales, y especialmente los logit, pueden ser utilizados para abordar la concordancia secuencial. De esta forma, los modelos log-lineales no sólo puede utilizarse para realizar el análisis secuencial, sino que pueden ser utilizados como una técnica de control de la calidad de los datos que se van a analizar.

Muy resumidamente, si A es una variable formada por las categorías de un sistema tomadas como antecedentes (ocurrencia de las categorías en el tiempo t), C las categorías de un sistema tomadas como consecuentes (y por tanto ocurrencia de las categorías en el tiempo $t + 1$), y R la variable registro (con las modalidades relativas por ejemplo a diferentes codificadores independientes), los modelos logit para evaluar la concordancia se pueden clasificar en tres tipos:

a) *Modelos tipo I*: las categorías consecuentes sólo dependen de ellas mismas. No hay secuencialidad, esto es, no es posible determinar patrones puesto que ninguna categoría consecuente es debida a una antecedente. Utilizando la notación propuesta por Goodman (1972) este modelo tendría la forma:

$$\Phi_{ik}^C = \beta^C \quad (1)$$

b) *Modelos tipo II*: los consecuentes dependen de los antecedentes (secuencialidad), pero no del registro (concordancia). Este modelo denominado de concordancia secuencial tendría la forma:

$$\Phi_{ik}^C = \beta^C + \beta_i^{CA} \quad (2)$$

Este tipo de modelos sugiere que las *categorías consecuentes* (tomadas como variable dependiente) son explicadas, además de por sí mismas, por las *categorías antecedentes*. Toda vez que los datos proceden de diversos registros y la variable *registro* no aparece en el modelo como independiente, puesto que sus efectos son irrelevantes para el ajuste, se puede concluir que las opiniones de los diferentes codificadores concuerdan.

Los modelos tipo II pueden además representarse gráficamente en un diagrama de paso, obteniéndose la siguiente figura:

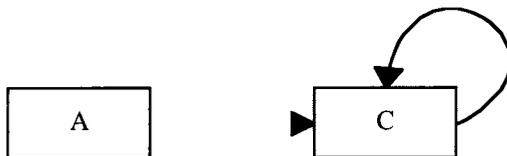


Figura 1. Diagrama de paso de los modelos de concordancia secuencial.

c) *Modelos tipo III*: los consecuentes dependen tanto de la interacción con los antecedentes (secuencialidad), como de la producida con los diferentes registros (discordancia). El modelo de *discordancia secuencial* tiene la forma:

$$\Phi_{ik}^C = \beta^C + \beta_i^{CA} + \beta_k^{CR} \quad (3)$$

Esta otra clase de modelos supone la inclusión del registro como variable explicativa, de manera que, según las diferentes modalidades del mismo, el resultado de las conexiones entre variables antecedentes y consecuentes difieren.

Si se representa este tipo de modelos mediante un diagrama de paso se obtiene la siguiente figura:

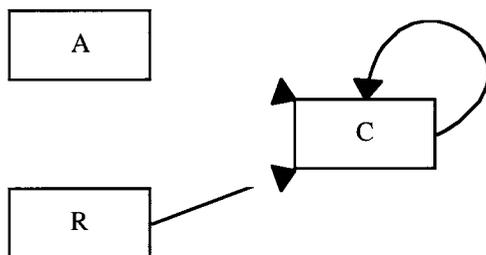


Figura 2. Diagrama de paso de los modelos de discordancia secuencial.

La selección de este modelo (tipo III) sugiere que los registros no son intercambiables.

d) *Modelos tipo IV* (modelo saturado): los consecuentes se hayan influidos por ellos mismos, por los antecedentes, por los registros y por la interacción de estos dos últimos.

$$\Phi_{ik}^C = \beta^C + \beta_i^{CA} + \beta_k^{CR} + \beta_{ik}^{CAR} \quad (4)$$

Tras la selección del modelo (que con parsimonia y buen ajuste represente adecuadamente los datos manteniendo una interpretación sustantiva), es posible medir la magnitud de la asociación entre las variables mediante la *concentración* y la *entropía* (Haberman, 1982).

Un aspecto de especial relevancia consiste en concretar la estructura del desacuerdo. Esto se puede realizar a partir de las *secuencias discordantes significativas* que se ponen de manifiesto estudiando los residuales entre las frecuencias observadas y esperadas. Una medida simple de discrepancia son los residuales estandarizados:

$$z = \frac{F - \hat{F}}{\sqrt{\hat{F}}} \quad (5)$$

Sin embargo, esta medida tiende a subestimar el valor real (Bakeman, Adamson y Strisik, 1989). Una mejor aproximación a z la ofrece el *residual ajustado* (Haberman, 1978: 77-79). El residual ajustado es además conceptualmente equivalente a la medida sugerida por Gottman (1980) y Allison y Liker (1982) para realizar el análisis secuencial.

REFERENCIAS BIBLIOGRÁFICAS

- AGRESTI, A. (1990): *Categorical data analysis*. New York: John Wiley & Sons.
- ALLISON, P. D. Y LIKER, J. K. (1982): Analyzing Sequential Categorical data on Dyadic Interaction: A Comment on Gottman. *Psychological Bulletin*, 91, (2), 393-403.
- ANGUERA, M. T. (1983): *Manual de Prácticas de Observación*. México: Trillas.
- ANGUERA, M. T. (1991). (Ed.): *Metodología observacional en investigación psicológica*. Barcelona: PPU.
- BAKEMAN, R. (1991): Prologue, en M. T. ANGUERA (Ed.): *Metodología observacional en investigación psicológica*. Barcelona: PPU, 13-17.
- BAKEMAN, R., ADAMSON, L. B. Y STRISIK, P. (1989): Lags and Logs: Statistical Approaches to Interaction, en M. H. BORNSTEIN y J. BRUNER (Eds.): *Interaction in Human Development* (241-260), Hillsdale, N. J.: Erlbaum.
- BAKEMAN, R. Y GOTTMAN, J. M. (1986): *Observación de la interacción: Introducción al análisis secuencial*, Madrid: Morata, 1989.
- GOODMAN, L. A. (1972): A general model for the analysis of surveys. *American Journal of Sociology*, 77, 1.035-1.086.
- GOTTMAN, J. M. (1979): *Marital interaction: Experimental investigation*. New York: Academic Press.
- GOTTMAN, J. M. (1980): On analyzing for sequential connection and assessing interobserver reliability for the sequential analysis of observational data. *Behavioral Assessment*, 2, 361-368.
- GOTTMAN, J. M. y ROY, A. K. (1990): *Sequential analysis. Guide for behavioral researchers*. Cambridge: Cambridge University Press.
- HABERMAN, S. J. (1982): Analysis of dispersion of multinomial responses. *Journal of the American Statistical Association*, 77, 568-580.
- SACKETT, G. P. (1979): The Lag Sequential Analysis of Contingency and Cyclicity in Behavioral Interaction Research, en J. D. OSOFSKY (Ed.): *Handbook of Infant Development*. New York: Wiley, 623-649.
- TANNER, R. y YOUNG, M. A. (1985): Modelling agreement among raters. *Journal of the American Statistical Association*, 80 (389), 175-180.
- TANNER, R. y YOUNG, M. A. (1986): Modeling ordinal scale disagreement. *Psychological Bulletin*, 98 (2), 408-415.
- TÓJAR, J. C. (1990): Momentos de una investigación observacional en la evaluación de una intervención en el aula. *Revista de Investigación Educativa*, 8, (16), 305-307.
- TÓJAR, J. C. (1992): *Concordancia del registro observacional en datos secuenciales. Investigación aplicada en el contexto del aula*. Málaga: Secretariado de Publicaciones de la Universidad de Málaga.
- WAMPOLD, B. E. y HOLLOWAY, E. L. (1983): «A note on interobserver reliability for sequential data». *Journal of Behavioral Assessment*, 5, (3), 755-765.

T.O.P.S.I.S.: UN MÉTODO DE DECISIÓN MULTIATRIBUTO COMO TÉCNICA EVALUATIVA

Rafael Galante Guille
Universidad de Málaga

El contenido de esta comunicación se enmarca dentro de las técnicas estadísticas del **Análisis de las Decisiones**, más concretamente los **Métodos de Decisión Multiatributo como instrumental de ayuda a la Evaluación Educativa**. Por las limitaciones inherentes a esta comunicación nos cofinaremos a la aplicación de una técnica concreta, cual es **Técnica de Ordenación de Preferencias por Similitud a la Solución Ideal**, en siglas **TOPSIS**, ya que supondremos que el decisor dispone de información cardinal sobre los atributos.

Los **Métodos de Decisión Multicriterios, MCDM**, se refieren a cuestiones de la toma de decisión en presencia de criterios múltiples, usualmente en conflictos. Estas situaciones problemáticas son muy usuales en la vida diaria.

Como por ejemplo: *en un contexto personal*, el trabajo que podemos escoger depende de muchos atributos, como el salario, el prestigio social, las oportunidades de progresar, las condiciones del trabajo (horario flexible, concurso de traslado, etc.), etc. Otra situación personal es cuando elegimos el compañero/a para formar una familia; la elección de nuestros estudios; la elección del vehículo para desplazarnos; etc.

En un contexto académico: la selección del rector de la universidad de la futura configuración de la universidad podría depender del número de escuelas, de facultades, escuelas superiores, de la distribución de las matrículas de los centros, de los planes de estudios, de los programas de investigación, etc. En este mismo contexto *la selección/evaluación de los candidatos a una plaza de profesor* puede ser una cuestión enfocable a través de la utilización de los Métodos de Decisión Multicriterios, como tendremos ocasión de estudiar en la presente comunicación.

En general, en cualquier contexto donde nos situemos, ya sea personal, académico, empresarial, público, etc. nos encontramos con situaciones de decisión que pueden ser auxiliadas con estos métodos.

Los problemas de decisión multicriterios son ampliamente diversos, sin embargo comparten las siguientes características comunes: *Atributos/Objetivos múltiples*, ya que el decisor debe generar atributos u objetivos relevantes para la situación problemática propuesta.

Conflicto entre criterios, ya que una mayor utilización o ponderación de un recurso o atributo supone una menor utilización o ponderación en otros.

Unidades incomparables, ya que cada objetivo/atributo suele venir medido en diferentes magnitudes de unidades.

Diseño/Selección, ya que las soluciones a estas situaciones problemáticas se entienden que son o **DISEÑAR** la 'mejor' alternativa o **SELECCIONAR** la 'mejor' de entre un conjunto finito de alternativas previamente especificado.

Así pues, los problemas de MCDM pueden ser clasificados, ampliamente, en dos categorías: los de diseño o Métodos de Decisión MultiObjetivos, **MODM**; y los de selección/evaluación o Métodos de Decisión Multiatributo, **MADM**.

La característica distintiva de los MADM es que, usualmente, tienen un número limitado, y generalmente pequeño, de alternativas predeterminadas. *Las alternativas están asociadas con un nivel de realización de los atributos, los cuales no necesariamente tienen que ser cuantificables. La selección final de la alternativa se hace con la ayuda de comparaciones inter e intra atributos. Las comparaciones pueden implicar compensaciones o negociaciones explícitas o implícitas.*

Históricamente, el esfuerzo de introducir el concepto de criterios múltiples en los procesos de toma de decisión comenzó en los años 60; sin embargo el estudio sobre los criterios múltiples tienen una larga tradición de los investigadores en diversas disciplinas, como por ejemplo: En **Teoría Matemática de la Decisión**: Teoría de la Utilidad Multiatributo, Probabilidad a Priori, Metodología Maximin. En **Teoría Económica**: Optimalidad de Pareto, Utilidad de Von Neumann-Morgenstern, Función de Bienestar Social, Análisis de Coste/Beneficio. En **Estadística**: Regresión Multivariante, Análisis Factorial. En **Psicometría**: Escalamiento Multidimensional, Medición Conjunta. Etc.

Los cuatro términos más usados en la literatura MCDM son atributos, objetivos, metas y criterios. No existe una definición universal para ellos, de hecho hay investigadores que los usan intercambiabilmente; sin embargo, en aras a la clarificación, precisaremos que un **un criterio es una medida de la efectividad**, constituye la base para la evaluación y emerge en forma de atributo u objetivo en la situación problemática en estudio. **Las metas son valores a priori o niveles de aspiración**, con frecuencia nos referimos a ellas como *restricciones* ya que son diseñadas para restringir o limitar el conjunto de alternativas. **Los atributos suministrarán un medio de evaluar los niveles de un objetivo**, cada alternativa puede ser caracterizada por un número de atributos. **Un objetivo generalmente indica la dirección de cambio deseada.**

Con el ánimo de expresar como la técnica TOPSIS puede servir de ayuda como herramienta de análisis en la evaluación educativa consideremos la siguiente situación problemática de decisión educativa:

Deseamos seleccionar/evaluar, en un contexto académico, a cinco candidatos a una cierta plaza de profesor, para lo cual se ha acordado que deben ser consideradas las siguientes características: X_1 (titulación académica que categorizaremos en idónea, adecuada o inadecuada); X_2 (experiencia docente que mediremos en años y que deberá ser corregida según sea idónea, adecuada o inadecuada con la plaza); X_3 (experiencia investigadora en grupos que mediremos en años); X_4 (experiencia empresarial que mediremos en años); X_5 (cursos conformes al perfil de la plaza que mediremos en número de horas); X_6 (publicaciones conformes con el perfil de la plaza que categorizaremos nominalmente en muy abundante, abundante, normal, escaza o muy escaza); X_7 (entrevista personal cuyo juicio podremos medir en idóneo, muy adecuado, adecuado, poco adecuado, inadecuado). Cualquier problema de MADM puede ser expresado en un lenguaje matricial.

Simularemos los niveles de valoración alcanzado por cada uno de los candidatos en cada atributo por la matriz de decisión expresada por:

	X_1	X_2	X_3	X_4	X_5	X_6	X_7
C_1	Idónea	6 años	8 años	0 años	100 h.	Normal	M. Adec.
C_2	Adec.	8 años	6 años	3 años	175 h.	Escasa	Idóneo
C_3	Inadec.	12 años	9 años	15 años	250 h.	Abund.	P. Adec.
C_4	Adec.	5 años	7 años	5 años	200 h.	M. Ab.	Adec.
C_5	Adec.	4 años	12 años	10 años	150 h.	M. Esc.	M. Adec.

Un primer problema con que nos enfrentamos es como **comparar atributos cuantitativos con atributos cualitativos** o borrosos y un segundo problema es que, en general, los atributos cuantitativos no están medidos en una misma unidad de medida con lo que se nos presenta **el problema de la homogeneidad de las unidades**. Estos son problemas de escalas de medidas. Una de las formas mas comunes para convertir un atributo cualitativo en escala de intervalo es utilizar la **escala bipolar**. Por ejemplo podemos seleccionar una escala de 10 puntos y comenzar con los puntos extremos, dando 10 puntos (0 puntos) al máximo (mínimo) valor que es prácticamente realizable; el punto medio será una base para la calibración ya que es el punto de rotura entre los valores que son favorables o mejor que la media y los que son desfavorables.

Así en X_1 (Idóneo=10, Adecuado=5, No adecuado=0); X_6 (Muy abundante=10, Abundante=7.5, Normal=5, Escasa=2.5, Muy escasa=0); X_7 (Idóneo=10, Muy adecuado=7.5, Adecuado=5, Poco adecuado=2.5, Nada adecuado=0).

Para resolver el segundo problema inherentes a la presencia de diferentes unidades en la matriz de decisión se recurre a la **normalización de los valores** de atributos a través de una transformación lineal de escala a una escala 0-1 por medio de restarle a cada valor el mínimo y a este resultado dividirlo por la diferencia entre el máximo y el mínimo valor, que podemos expresar algebraicamente por la expresión:

$r_{ij} = (x_{ij} - \text{Min}_i\{x_{ij}\}) / (\text{Max}_i\{x_{ij}\} - \text{Min}_i\{x_{ij}\})$; donde x_{ij} expresa el valor de la i-ésima alternativa en el j-ésimo atributo.

	X_1	X_2	X_3	X_4	X_5	X_6	X_7
C_1	1	0.25	1/3	0	0	0.5	2/3
C_2	0.5	0.5	0	0.2	0.5	0.25	1
C_3	0	1	0.5	1	1	0.75	0
C_4	0.5	0.125	1/6	1/3	2/3	1	1/3
C_5	0.5	0	1	2/3	1/3	0	2/3

En MADM una solución óptima, solución superior, **solución ideal** o punto de utopía es una alternativa cuyo producto cartesiano está compuesto de los valores más preferibles para cada uno de los atributos dados de la matriz de decisión.

Utilizaremos una **modelización compensatoria** que permite intercambio entre los atributos y dentro de éstos una **modelización de compromiso** y más precisamente la técnica TOPSIS que selecciona aquella alternativa cuya distancia es la menor a la solución ideal y la mayor a la solución ideal-negativa, supondremos que el decisor dispone de información de tipo cardinal o ponderación lineal convexa w_j sobre los atributos, $w_j \geq 0$ y $\sum_j w_j = 1$. En nuestro caso supondremos el vector de ponderación: (0.15, 0.2, 0.2, 0.15, 0.1, 0.15, 0.05). Construimos la matriz de decisión normalizada ponderada, $v_{ij} = w_j r_{ij}$.

	X_1	X_2	X_3	X_4	X_5	X_6	X_7
C_1	0.15	0.05	1/15	0	0	0.075	1/30
C_2	0.075	0.1	0	0.03	0.05	0.0375	0.05
C_3	0	0.2	0.1	0.15	0.1	0.1125	0
C_4	0.075	0.025	1/30	0.05	1/15	0.15	1/60
C_5	0.075	0	0.2	0.1	1/30	0	1/30

Determinamos la solución ideal A^+ y la solución ideal-negativa A^- :

$$v^+ = (\text{Máx}_i \{v_{ij}\}) = (0.15, 0.2, 0.2, 0.15, 0.1, 0.15, 0.05)$$

$$v^- = (\text{Min}_i \{v_{ij}\}) = (0, 0, 0, 0, 0, 0, 0)$$

Calculamos la medida de separación de cada alternativa C_i con la solución ideal S_{i^+} y con la solución ideal negativa S_{i^-} .

$$S_{i^+} = \sqrt{\sum_{j=1}^n (v_{ij} - v_j^+)^2} \quad S_{i^-} = \sqrt{\sum_{j=1}^n (v_{ij} - v_j^-)^2}$$

$$S_{1^+} = .28050054499848, S_{1^-} = .19021186988315$$

$$S_{2^+} = .29185827039849, S_{2^-} = .15143067720908$$

$$S_{3^+} = .19080421903092, S_{3^-} = .30847406698132$$

$$S_{4^+} = .27613402579182, S_{4^-} = .19257033312533$$

$$S_{5^+} = .27449448496787, S_{5^-} = .2405144947353$$

Calculamos la proximidad relativa del i -ésimo candidato a la solución ideal, $C(i)$, por la expresión $C(i) = S_{i^-} / (S_{i^+} + S_{i^-})$. Así una alternativa es más próxima a la solución ideal cuanto más próximo a la unidad sea su $C(i)$. Obtenemos:

$$C(1) = .40409359062897, C(2) = .34160715719702$$

$$C(3) = .61783994141848, C(4) = .41085671481747$$

$$C(5) = .46701029343978$$

Así el orden de selección de los candidatos es en primer lugar el tercer candidato, en segundo lugar el quinto candidato, en tercer lugar el cuarto candidato, en cuarto lugar el primer candidato y en último lugar el segundo candidato.

Como conclusión hemos podido observar como una técnica de decisión multiatributo, cual es la técnica de ordenación de preferencias por similitud a la solución ideal desarrollada por Yoon y Hwang (1980) sirve como herramienta de ayuda en los procesos de evaluación educativa siempre que el decisor disponga de información de tipo cardinal sobre los atributos a través de los cuales valorará cada una de las posibles alternativas.

REFERENCIAS BIBLIOGRÁFICAS

- HWANG, C. L. & YOON, K. (1980): «Principles for Evaluation of Air-Conditioning System». *A Progress Report to U.S. Office of Energy from Industrial Engineering*. Kansas State University.
- YOON, K. (1980): «System Selection by Multiple Attribute Decision Making». *Ph. D. Dissertation*. Kansas State University.

ANÁLISIS DEL FUNCIONAMIENTO DIFERENCIAL DE ÍTEMS DE RESPUESTA GRADUADA: UNA APROXIMACIÓN METODOLÓGICA

F. Javier Murillo Torrecilla

Centro de Investigación, Documentación y Evaluación (M.E.C.). Servicio de Investigación

María Castro Morera

Universidad Complutense de Madrid

En esta comunicación se aborda el estudio, desde un punto de vista metodológico y en base a la I.R.T., del funcionamiento diferencial de un tipo particular de ítems: los ítems de respuesta graduada. Así, se proponen tres procedimientos para calcular el «grado de sesgo» de ítems respuesta dicotómica, se presenta un modelo para estudiar ítems de respuesta graduada y, combinando ambas aproximaciones, se proponen tres métodos para calcular el grado de sesgo en ítems de respuesta graduada.

FUNCIONAMIENTO DIFERENCIAL DE LOS ÍTEMS

Intuitivamente se puede decir que un ítem funciona diferencialmente para dos grupos si dos sujetos de distintos grupos con el mismo nivel de habilidad tienen diferente rendimiento en ese ítem. Esta idea se puede traducir desde la perspectiva de la I.R.T. en la siguiente proposición: un ítem se comporta igual con dos grupos si las curvas características de ese ítem (C.C.I.) son idénticas para ambos grupos (Holland y Wainer, 1993). Dos definiciones equivalentes pueden ser:

- 1) Un ítem se comporta igual para dos grupos si el área comprendida entre las dos curvas características de los ítems es cero para todo entorno de cualquier nivel de habilidad (q) (Hambleton y Swaminathan, 1985).
- 2) Un ítem funciona igual para dos grupos si, para todo par de sujetos pertenecientes a distintos grupos y con el mismo nivel de habilidad, la probabilidad de acierto de ese ítem es idéntica (Pine, 1977; Lord, 1980; Hambleton y Swaminathan, 1985).

De la primera definición se deduce el **procedimiento de las áreas** para conocer el D.I.F., dado que el área comprendida entre las dos C.C.Is. se puede hallar calculando la integral de la diferencia de las funciones de probabilidad. Así, se establece el «grado de sesgo» a través de la siguiente fórmula:

$$Sesgo = \int_{-\infty}^{\infty} [P_1^1(\theta) - P_1^2(\theta)] d(\theta)$$

donde $P_1^1(q)$ y $P_1^2(q)$ son las funciones que definen las curvas de cada grupo.

Esta integral definida es fácil de calcular para el Modelo logístico de un parámetro. En ese caso se obtiene:

$$sesgo = \int_{-\infty}^{\infty} \left[\frac{e^{(\theta-b)}}{1+e^{(\theta-b)}} - \frac{e^{(\theta-b')}}{1+e^{(\theta-b')}} \right] d(\theta) = b' - b$$

Este resultado, fácil y práctico, ofrece la posibilidad de calcular una aproximación al grado de sesgo estimando únicamente los parámetros b de cada grupo.

Para otro tipo de modelos, dada la dificultad que supone resolver esa integral, se puede realizar una aproximación numérica de la misma. Aunque la forma más habitual de hacerlo es mediante la fórmula de los rectángulos (Linn y cols., 1981), se obtiene una mejor **aproximación mediante la fórmula de Simpson**. Así:

$$sesgo = \int_a^b P_i^*(\theta) d(\theta) \approx \frac{b-a}{6m} [y_0 + y_{2m} + 2(y_2 + y_4 + \dots + y_{2m-2}) + 4(y_1 + y_3 + \dots + y_{2m-1})]$$

donde $2m$ es el número de intervalos, y_0, y_1, \dots, y_{2m} son los valores de la función diferencia $P_i^*(q)$ en los puntos q_0, \dots, q_{2m} , a es el máximo valor para el cual las C.C.Is. de cada grupo son simultáneamente cero y b es el mínimo valor donde las dos C.C.Is. toman el valor de uno.

De la segunda definición propuesta se extrae el procedimiento que se denominará **de las probabilidades empíricas**. Consiste en comparar las probabilidades empíricas de acierto para cada grupo de sujetos con niveles de habilidad similares. Así, se calibra el test para toda la población y se halla el nivel de habilidad para cada sujeto. En un segundo paso, se agrupan los niveles de habilidad en intervalos, (por ejemplo, 12 intervalos entre $q=-3$ y $q=3$) se hallan las probabilidades de acierto de cada intervalo de habilidad separadamente para cada grupo (y_i^j) y, con estas puntuaciones, se calcula el grado de sesgo mediante la siguiente fórmula:

$$sesgo = \frac{b-a}{n} [(y_1^1 - y_1^2) + (y_2^1 - y_2^2) + \dots + (y_n^1 - y_n^2)] = \frac{b-a}{n} \sum_{i=1}^n (y_i^1 - y_i^2)$$

donde y_i^1 e y_i^2 son las probabilidades empíricas en cada intervalo i por los grupos 1 y 2.

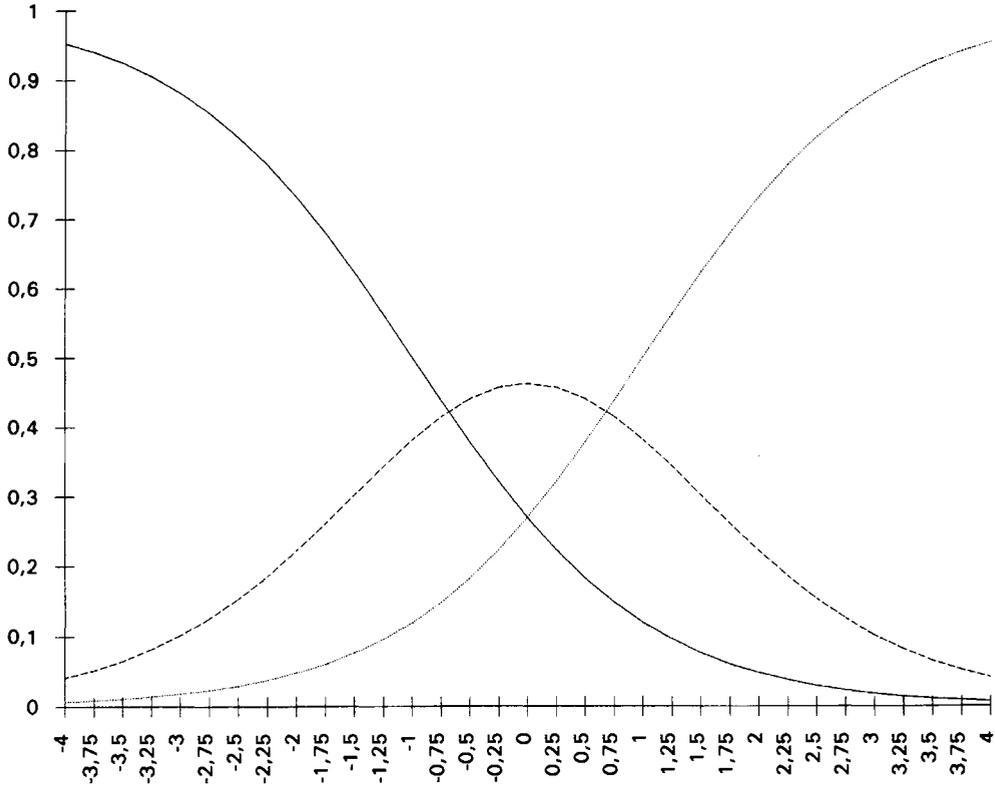
MODELO PARA ÍTEMS DE RESPUESTA GRADUADA

Los modelos tradicionales de I.R.T. están elaborados para ítems respuesta dicotómicas. Sin embargo, existen muchas pruebas donde es posible encontrar ítems con otras alternativas; por ejemplo, con las alternativas de correcto, parcialmente correcto e incorrecto. Para esos ítems, llamados de respuesta graduada, Samejima (1969) propuso el denominado **modelo diferencial**. Según el mismo, cada ítem está definido por tantas curvas como posibles respuestas del mismo. Cada una de las curvas está definida por la siguiente fórmula:

$$P_i(\theta) = P_i^*(\theta) - P_{i+1}^*(\theta)$$

para $i=1, \dots, n-1$, donde $P^*(q)$ es la regresión de la puntuación de un ítem dicotómico sobre la habilidad, cuando todas las categorías de respuesta menores que i son puntuadas como 0, y aquellas mayores o iguales a i se puntúan 1. Esta propuesta sirve para cualquier modelo de respuesta al ítem. En la siguiente gráfica se pueden observar las curvas características de los ítems que representan cada una de las tres posibles respuestas de este ítem:

CCIs de un ítem con el Modelo de Respuesta Graduada



Para el Modelo Logístico de dos parámetros las curvas características para ítems de respuesta graduada están definidas por:

$$P_0(\theta) = \frac{1}{1 + e^{a_1(\theta - b_1)}}$$

$$P_i(\theta) = \frac{e^{a_i(\theta - b_i)} - e^{a_{(i+1)}(\theta - b_{(i+1)})}}{1 + e^{a_i(\theta - b_i)} + e^{a_{(i+1)}(\theta - b_{(i+1)})} + e^{a_i(\theta - b_i) + a_{(i+1)}(\theta - b_{(i+1)})}}$$

para $i = 1, \dots, n-1$.

$$P_n(\theta) = \frac{e^{a_n(\theta - b_n)}}{1 + e^{a_n(\theta - b_n)}}$$

FUNCIONAMIENTO DIFERENCIAL DE ÍTEMS DE RESPUESTA GRADUADA

En este tercer bloque se aplicarán los tres procedimientos de detección del D.I.F.: de las áreas, de la aproximación numérica a través de la fórmula de Simpson y de las probabilidades empíricas al modelo de ítems de respuesta graduada descrito.

Dado que los ítems de respuesta graduada no tienen una sola curva característica sino tantas como posibles respuestas del ítem (n), los procedimientos de detección de D.I.F. deberán aplicarse a cada una de las curvas que definen el ítem, o, con más exactitud, a n-1 de las curvas que lo definen. Ello tiene un sentido lógico, dado que puede darse el caso de un ítem que se comporte de manera diferente para una o varias de las alternativas pero no para el ítem tomado conjuntamente.

Así, si eliminamos la curva $P_0(q)$, y aplicando el **procedimiento de las áreas** habría que calcular:

$$Sesgo^{[j]} = \int_{-\infty}^{\infty} [P_i^{[j] (1)}(\theta) - P_i^{[j] (2)}(\theta)] d(\theta)$$

para $j = 1, \dots, n$.

Sustituyendo:

$$Sesgo^{[j]} = \int_{-\infty}^{\infty} [(P_i^{*[j] (1)}(\theta) - P_i^{*[j+1] (1)}(\theta)) - (P_i^{*[j] (2)}(\theta) - P_i^{*[j+1] (2)}(\theta))] d(\theta)$$

para $j = 1, \dots, n-1$, y

$$Sesgo^{[n]} = \int_{-\infty}^{\infty} [P_i^{*[n] (1)}(\theta) - P_i^{*[n] (2)}(\theta)] d(\theta)$$

Estas fórmulas son generales para cualquier modelo de I.R.T.

Si se utiliza el modelo de Rasch, la integral tiene fácil solución:

$$Sesgo^{[j]} = (b_i^{*([j](1))} - b_i^{*([j+1](1))}) - (b_i^{*([j](2))} - b_i^{*([j+1](2))}) \text{ para } j=1, \dots, n-1$$

$$Sesgo^{[n]} = b_i^{*([n](1))} - b_i^{*([n](2))}$$

El otro procedimiento a analizar es el que se ha denominado **procedimiento de las áreas a través de la fórmula de Simpson**. Según el mismo, el sesgo para cada una de las curvas que definen el ítem se calcula de la siguiente manera:

$$Sesgo^{[j]} = \frac{b-a}{6m} [y_0^{[j]} + y_{2m}^{[j]} + 2(y_2^{[j]} + y_4^{[j]} + \dots + y_{2m-2}^{[j]}) + 4(y_1^{[j]} + y_3^{[j]} + \dots + y_{2m-1}^{[j]})]$$

donde [j] corresponde a cada una de las curvas diferenciales y 2m es el número de intervalos.

Por último, el **procedimiento de las probabilidades empíricas** se resuelve mediante la siguiente fórmula:

$$Sesgo^{[j]} = \frac{b-a}{n} \sum_{i=1}^n (y_i^{[j]1} - y_i^{[j]2})$$

donde [j] corresponde a cada una de las curvas diferenciales.

CONCLUSIONES

Las aportaciones de la comunicación presentada se pueden resumir en los siguientes puntos:

1. Se han propuesto tres **procedimientos para el cálculo del funcionamiento diferencial de los ítems**.

- a) En primer lugar, mediante el procedimiento de las áreas, se ha buscado una solución para el Modelo de Rasch. Esta solución está caracterizada por la extrema sencillez que supone su puesta en marcha, puesto que la estimación del «grado de sesgo» depende exclusivamente de los parámetros de dificultad, sin embargo, tiene el inconveniente de perder mucha información si el modelo no se ajusta a los datos.
 - b) En segundo lugar, se ha realizado una mejor estimación del área comprendida entre las curvas características de los ítems mediante la fórmula de Simpson. Con este cálculo se ofrece una aproximación bastante exacta al «grado de sesgo», aunque sigue teniendo el inconveniente de partir de las probabilidades estimadas a través del modelo.
 - c) Por último, se ha propuesto el procedimiento denominado de las probabilidades empíricas, que tiene la ventaja de trabajar con las puntuaciones directas, por lo que aporta información muy ajustada de la realidad. Sin embargo, tiene como inconveniente que es necesario un importante número de sujetos para poder utilizar la propuesta.
2. Se ha desarrollado el **modelo diferencial** de Samejima (1969) para ítems de respuesta graduada.
3. Se han aplicado los tres procedimientos de detección del D.I.F. al modelo de ítems de respuesta graduada, obteniéndose tres procedimientos novedosos para estudiar el **funcionamiento diferencial de los ítems de respuesta graduada** de una manera sencilla y práctica.

REFERENCIAS BIBLIOGRÁFICAS

- HAMBLETON, R. K. y SWAMINATHAN, H. (1985): *Item response theory. Principles and applications*. Boston: Kluwer-Nijhoff Pub.
- HOLLAND, P. W. y WAINER, H. (1993): *Differential ítem functioning*. Hillsdale, N.J.: LEA.
- LINN, R. L., LEVINE, M. V., HASTINGS, C. N. y WARDROP, J. L. (1981): An investigation of item bias in a test of reading comprehension. *Applied Psychological Measurement*, 5, 159-173.
- LORD, F. M. (1980): *Applications of ítem response theory to practical testing problems*. Hillsdale, N.J.: LEA.
- PINE, S. M. (1977): Applications of item response theory to the problem of test bias. En D. J. WEISS (ed.): *Applications of computerized adaptive testing*. Research Report 77-1. Minneapolis, University of Minnesota, Psychometric Methods Program, Department of Psychology.
- SAMEJIMA, R. (1969): Estimation of latent ability using a response pattern of graded scores. *Psychometric Monograph*, 17. Psychometric Society.

PROCEDIMIENTO PARA LA DETERMINACIÓN ARBITRARIA DE LOS VALORES DE UNA MATRIZ DE CORRELACIONES

José Luis Gaviria Soto
Universidad Complutense de Madrid

1. INTRODUCCIÓN

En este artículo se presenta un método para determinar los valores máximo y mínimo que pueden tomar los elementos de una matriz de intercorrelaciones. La determinación de estos valores es muy útil cuando se desea generar un conjunto de variables con una cierta estructura correlacional o causal. Estos conjuntos de variables así generados tienen dos tipos de aplicaciones. El primero, para llevar a cabo simulaciones o experimentos Montecarlo. El segundo, en la docencia de distintos tipos de técnicas de análisis de datos.

En el primer caso, la determinación apriorística de una estructura latente a un conjunto de variables puede hacerse tanto imponiendo una determinada matriz de intercorrelaciones, como explicitando una determinada estructura de variables latentes. Pero también en esta segunda posibilidad es necesario determinar en algún momento la estructura de una matriz de intercorrelaciones entre variables latentes. Ciertamente puede ocurrir que un pequeño número de variables latentes expliquen un número mayor de variables simuladas, pero tan pronto como estas últimas crezcan en número, nos veremos abocados a, o bien mantener una estructura latente excesivamente simple, o bien a encontrarnos con que las matrices de intercorrelaciones que especificamos presentan inconsistencias que las inhabilitan.

En un programa como EQS (Bentler, 1989) se da, por ejemplo, la posibilidad de generar datos por simulación por uno de dos procedimientos: especificando un modelo causal determinado, o bien a partir de una matriz de intercorrelaciones. Si la matriz que se especifica, o las covarianzas entre las variables independientes son inconsistentes, la ejecución del programa dará como resultado el intento de obtener una correlación múltiple al cuadrado mayor que 1. Con el procedimiento que aquí se presenta se evita en cualquier caso el peligro de esas inconsistencias. Este procedimiento es por otra parte fácilmente programable en cualquier lenguaje de alto nivel, con lo que puede obtenerse una ayuda interactiva para la generación de la matriz objetivo.

2. LÍMITES MÁXIMO Y MÍNIMO DE UNA CORRELACIÓN ENTRE DOS VARIABLES DE UN CONJUNTO INTERCORRELACIONADO

Si tenemos una matriz de intercorrelaciones de por ejemplo cinco filas y cinco columnas, la determinación de los elementos de cada fila implica a los elementos de otras filas de esa matriz. Por ejemplo si queremos determinar la correlación de la variable 3 con la 1 y la 2, ocurrirá que el límite superior de la primera correlación que determinemos es 1, es decir, tenemos libertad absoluta para

asignar ese valor. $\varphi_{31} \leq 1$. Pero para determinar la segunda, φ_{32} , el límite será una función de la correlación existente entre las variables 1 y 2, y 3 y 1. $\varphi_{32} = (\varphi_{12}, \varphi_{31})$. En general,

$$\varphi_{ij} = f(\varphi_{ik}, \varphi_{im}) \quad \forall k < j, \quad i = 1 \dots j-1, \quad m = 2 \dots j$$

Es decir, el límite máximo que puede tomar una correlación es función de todos los elementos anteriores de su fila, y de todos los elementos de las filas iguales o inferiores en orden al orden de la columna a la que pertenece.

3. INVERSA DE UNA MATRIZ EN FUNCIÓN DE SUS PARTICIONES

Una matriz particionada, es una matriz cuyos elementos son a su vez matrices de orden inferior. Así,

$$A = \begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}$$

Más adelante nos será de utilidad el poder expresar la inversa de una matriz en función de sus particiones. $A^{-1} = f(A_{11}, A_{12}, A_{21}, A_{22})$. La inversa de A, A^{-1} , será una matriz B, que expresada en particiones será

$$\begin{array}{c|c} B_{11} & B_{12} \\ \hline B_{21} & B_{22} \end{array}$$

y se cumplirá que $\begin{array}{c|c} I & 0 \\ \hline 0 & I \end{array}$. Por lo tanto tenemos un conjunto de ecuaciones

$A_{11} B_{11} + A_{12} B_{21} = I$
$A_{11} B_{12} + A_{12} B_{22} = 0$
$A_{21} B_{11} + A_{22} B_{21} = 0$
$A_{21} B_{12} + A_{22} B_{22} = I$

Resolviendo este sistema de ecuaciones, (Véase HARRIS, (1985), pp. 370-373) obtenemos

$$A^{-1} = \frac{A_{11}^{-1} + (A_{11}^{-1} A_{12} (A_{22} - A_{21} A_{11}^{-1} A_{12})^{-1} A_{21} A_{11}^{-1})}{-(A_{22} - A_{21} A_{11}^{-1} A_{12})^{-1} A_{21} A_{11}^{-1}} \quad \left| \quad \frac{-A_{11}^{-1} A_{12} (A_{22} - A_{21} A_{11}^{-1} A_{12})^{-1}}{(A_{22} - A_{21} A_{11}^{-1} A_{12})^{-1}} \right.$$

En el caso particular de una matriz de correlaciones que es simétrica, haciendo

$$R_j = \frac{R_{j-1}}{V'_{j-1}} \quad \left| \quad \frac{V_{j-1}}{1} \right. \quad \text{siendo } A_{11}=R_{j-1}, \quad A_{12}=V_{j-1}, \quad A_{21}=V'_{j-1}, \quad A_{22}=1 \text{ y } A=R_j \text{ tendremos}$$

$$R_j^{-1} = \frac{R_{j-1}^{-1} + (R_{j-1}^{-1} V_{j-1} (1 - V'_{j-1} R_{j-1}^{-1} V_{j-1})^{-1} V'_{j-1} R_{j-1}^{-1})}{-(1 - V'_{j-1} R_{j-1}^{-1} V_{j-1})^{-1} V'_{j-1} R_{j-1}^{-1}} \quad \left| \quad \frac{-R_{j-1}^{-1} V_{j-1} (1 - V'_{j-1} R_{j-1}^{-1} V_{j-1})^{-1}}{(1 - V'_{j-1} R_{j-1}^{-1} V_{j-1})^{-1}} \right.$$

y haciendo $1 - V'_{j-1} R_{j-1}^{-1} V_{j-1} = d$

$$R_j^{-1} = \frac{R_{j-1}^{-1} + \frac{R_{j-1}^{-1}V_{j-1}V_{j-1}'R_{j-1}^{-1}}{d}}{\frac{V_{j-1}'R_{j-1}^{-1}}{d}} \quad \left| \quad \begin{array}{c} -\frac{R_{j-1}^{-1}V_{j-1}}{d} \\ \frac{1}{d} \end{array} \right.$$

4. CONDICIÓN QUE DEBE CUMPLIR UNA MATRIZ DE CORRELACIONES

Supongamos que tenemos definidas un conjunto de variables cuya matriz de intercorrelaciones es R_n . Suponemos también que esa matriz tiene inversa, R_n^{-1} . Si queremos añadir una variable más a ese conjunto inicial, tendremos entonces una matriz de intercorrelaciones R_{n+1} que tendrá una fila y una columna más que R_n . Los elementos de esas fila y columna serán las correlaciones de la nueva variable con todas las anteriores. A la matriz de correlaciones entre las n primeras variables podemos denominarla R . El vector formado por los elementos $\phi_{1n+1}, \phi_{2n+1}, \dots, \phi_{nn+1}$ lo denominamos vector V . Por lo tanto la correlación múltiple de la variable $n+1$ con las n anteriores vendrá dada por

$$\rho_{n+1.1,2,\dots,n}^2 = V'R^{-1}V.$$

Lógicamente el valor máximo que puede tomar la correlación múltiple al cuadrado es 1. Por lo tanto nuestro problema se reduce a encontrar un conjunto de valores de los elementos de la matriz R tales que $V'R^{-1}V \leq 1$.

5. LÓGICA DEL PROCESO

Añadir una variable más a un conjunto inicial supone generar un vector de correlaciones de esa nueva variable con cada una de las n variables iniciales. De ese nuevo vector, el primer elemento, puede adoptar cualquier valor entre 1 y -1 , es totalmente libre. Pero todos los demás elementos irán dependiendo de los anteriores, como se señala en el primer apartado de este artículo.

La determinación de los valores máximo y mínimo de cada uno de esos elementos ij , siendo $i > j^1$ y $j > 1^2$, es lo que se desarrolla en este apartado. Para una mejor comprensión de la notación y las definiciones que utilizamos en esta sección, vamos a hacer una presentación de la lógica empleada. Como ejemplo, supongamos que la matriz que queremos construir, (fig. 6.1) tiene un total de 6 filas y 6 columnas. El proceso de construcción implica que vamos generando las correlaciones por columnas. La generación es de tal forma que si tenemos ya definidas las correlaciones señaladas en la matriz de correlaciones en el recuadro superior izquierdo, las correlaciones de los recuadros de las tres líneas siguientes son independientes entre sí, es decir, al generar las correlaciones del vector V_3^4 , se deberá cumplir la condición de que $V_{43}'R_3^{-1}V_{43} \leq 1$, y lo mismo con cada uno de los vectores, es decir, $V_{53}'R_3^{-1}V_{53} \leq 1$ y $V_{63}'R_3^{-1}V_{63} \leq 1$. Como puede observarse para generar esos vectores, (de hecho sólo el último elemento de cada vector, ya que los otros se han generado en una fase anterior), es preciso invertir la matriz R_3 . En la siguiente fase, cuando se trate de generar los elementos de la columna 4, (2 elementos que están por debajo de la diagonal), habrá que invertir la matriz R_4 . Dado que hay que invertir varias

1 Esta condición se refiere a que sólo hace falta calcular el triángulo inferior, ya que la matriz es simétrica.

2 Esta otra condición hace referencia al hecho de que todas las correlaciones de la primera columna no están determinadas por ninguna otra, luego sus valores máximo y mínimo son $+1$ y -1 .

R_3	1	Φ_{12}	Φ_{13}	Φ_{14}	Φ_{15}	Φ_{16}
	Φ_{21}	1	Φ_{23}	Φ_{24}	Φ_{25}	Φ_{26}
	Φ_{31}	Φ_{32}	1	Φ_{34}	Φ_{35}	Φ_{36}
V'_{43}	Φ_{41}	Φ_{42}	Φ_{43}	1	Φ_{45}	Φ_{46}
V'_{53}	Φ_{51}	Φ_{52}	Φ_{53}	Φ_{54}	1	Φ_{56}
V'_{63}	Φ_{61}	Φ_{62}	Φ_{63}	Φ_{64}	Φ_{65}	1

matrices en todo el proceso, y que tales matrices se forman simplemente como yuxtaposición de varias particiones, es de interés poner la inversa de cada una de esas matrices en función de la inversa de orden anterior, como se vio en el apartado 2. Es decir,

$$R_4 = \frac{R_3}{V'_{43}} \left| \begin{array}{c} V_{43} \\ 1 \end{array} \right. \text{ y a su vez } R_3 = \frac{R_2}{V'_{32}} \left| \begin{array}{c} V_{32} \\ 1 \end{array} \right., \text{ luego en general } R_j = \frac{R_{j-1}}{V'_{j-1}} \left| \begin{array}{c} V_{j-1} \\ 1 \end{array} \right. \text{ y por lo tanto}$$

la inversa queda según la expresión de 3.1 y 3.2.

6. VALORES MÁXIMO Y MÍNIMO DE UNA CORRELACIÓN

Definimos $M'_j = \Phi_{j1}, \Phi_{j2}, \dots, \Phi_{j-1}$, y $V'_j = \Phi_{j1}, \Phi_{j2}, \dots, \Phi_{j-1} | \Phi_j = M'_j | \Phi_j$

por lo que el producto que debe cumplir la condición que se señalaba en el 4 será

$$\begin{aligned}
 V'_j R_j^{-1} V_j &= M'_j | \Phi_j \left| \begin{array}{c} \frac{dR_{j-1}^{-1} + R_{j-1}^{-1} V_{j-1} V'_{j-1} R_{j-1}^{-1}}{d} \\ -\frac{V'_{j-1} R_{j-1}^{-1}}{d} \end{array} \right| \left| \begin{array}{c} -\frac{R_{j-1}^{-1} V_{j-1}}{d} \\ \frac{1}{d} \end{array} \right. \frac{M_j}{\Phi_j} = \\
 &= \frac{\Phi_j^2 - \Phi_j M'_j R_{j-1}^{-1} V_{j-1} - \Phi_j V'_{j-1} R_{j-1}^{-1} M_j + M'_j R_{j-1}^{-1} V_{j-1} V'_{j-1} R_{j-1}^{-1} M_j + d(M'_j R_{j-1}^{-1} M_j)}{d} = \\
 &= \frac{\Phi_j^2 - 2\Phi_j V'_{j-1} R_{j-1}^{-1} M_j + (M'_j R_{j-1}^{-1} V_{j-1})^2 + d(M'_j R_{j-1}^{-1} M_j)}{d}
 \end{aligned}$$

Como se ha de cumplir la condición señalada en 4, tendremos

$$\frac{\Phi_j^2 - 2\Phi_j V'_{j-1} R_{j-1}^{-1} M_j + (M'_j R_{j-1}^{-1} V_{j-1})^2 + d(M'_j R_{j-1}^{-1} M_j)}{d} \leq 1$$

que es una ecuación de segundo grado. Resolviendo la misma obtenemos

$$\Phi_{ij} = M'_j R_{j-1}^{-1} V_{j-1} \pm \sqrt{(1 - V'_{j-1} R_{j-1}^{-1} V_{j-1})(1 - M'_j R_{j-1}^{-1} M'_j)}$$

Estos son los límites superior e inferior de los valores que puede tener la correlación ij, dadas todas las anteriores.

7. INTERPRETACIÓN DEL RESULTADO

En la expresión que hemos obtenido, podemos multiplicar por $R_{j-1}^{-1} R_{j-1} = I$, con lo que

$$\begin{aligned} \Phi_{ij} &= M'_j R_{j-1}^{-1} R_{j-1} R_{j-1}^{-1} V_{j-1} \pm \sqrt{(1 - V'_{j-1} R_{j-1}^{-1} V_{j-1})(1 - M'_j R_{j-1}^{-1} M'_j)} = \\ &= B'_{ij-1} R_{j-1} B_{j-1} \pm \sqrt{(1 - \rho_{jj-1}^2)(1 - \rho_{ij-1}^2)} \end{aligned}$$

donde

$B'_{ij-1} = \beta_{i,1}, \beta_{i,2}, \dots, \beta_{i,j-1}$, vector de coeficientes de regresión múltiple, siendo i la variable dependiente y 1...j-1 las variables independientes;

$B'_{ij-1} = \beta_{j,1}, \beta_{j,2}, \dots, \beta_{j,j-1}$, vector de coeficientes de regresión múltiple, siendo j la variable dependiente y 1...j-1 las variables independientes;

$$R_{j-1} = \begin{matrix} \Phi_{11} & \dots & \Phi_{1j-1} \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \Phi_{j-11} & \dots & \Phi_{j-1j-1} \end{matrix} \text{ matriz de intercorrelaciones de las variables } 1\dots j-1;$$

ρ_{jj-1}^2 , correlación múltiple al cuadrado de la variable j con las variables 1...j-1;

ρ_{ij-1}^2 , correlación múltiple al cuadrado de la variable i con las variables 1...j-1.

Geoméricamente el resultado obtenido puede entenderse de la siguiente forma: las j-1 primeras variables son la base de un espacio. Las variables i y j añaden dos dimensiones más a ese espacio inicial. La matriz R_{j-1} refleja las posiciones relativas de cada una de las dimensiones de ese espacio, ya que sus elementos Φ_k son los cosenos de los ángulos formados por las variables l y k. Los vectores B son los cosenos de los ángulos formados por las variables i y j con las j-1 dimensiones de la base del espacio de j-1 dimensión. El producto $B'_{ij-1} R_{j-1} B_{j-1}$ es el valor del coseno del ángulo que forma la variable i con la variable j, cuando el subespacio formado por el vector i y las j-1 variables es ortogonal al subespacio formado por el vector j y las j-1 variables. Este concepto es muy fácil de entender si nos ceñimos a sólo tres variables. Supongamos que determinamos la correlación que deseamos entre dos variables, 1 y 2. Esa correlación, suponiendo vectores de módulo 1, es el coseno del ángulo α . Podemos seguidamente determinar el ángulo que deseamos entre los vectores 1 y 3. Es el ángulo β . Los vectores 1 y 2 forman un subespacio de dos dimensiones en este caso, y los vectores 1 y tres otro subespacio distinto de dos dimensiones. Si el vector tres revolucionara alrededor del vector 1 con el ángulo constante, obtendríamos un cono de revolución. Todos los vectores incluidos en ese cono nos dan, con el vector 2 todos los valores de los ángulos que pueden darse entre el vector 2 y el 3.

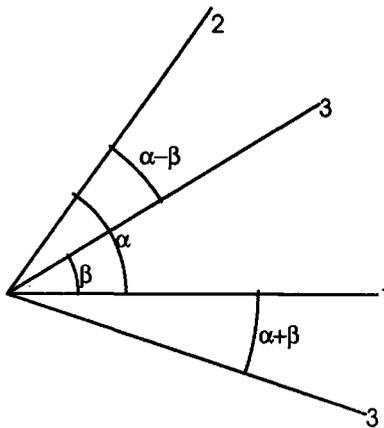
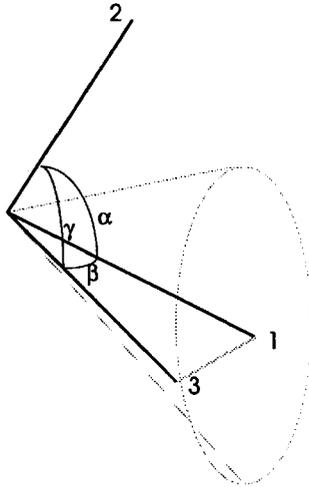
La expresión que estamos analizando nos da el ángulo que forman el vector 2 y el 3 cuando los subespacios de dos dimensiones determinados por los vectores 1 y 3, y 2 y 3 respectivamente, son ortogonales entre sí. En la figura están dibujados precisamente los tres vectores en esa posición. Puede apreciarse que el plano 1-3 es perpendicular al plano 1-2. Nuestra expresión nos da precisamente el coseno del ángulo que forman en esa posición los vectores 2 y 3.

Esta visión geométrica del problema ayuda a entenderlo correctamente. Precisamente los valores máximo y mínimo de la correlación entre los vectores 2 y 3 se alcanzan cuando ambos están situados junto con el vector 1 en el mismo plano, es decir, en el mismo subespacio de dos dimensiones.

En ese caso el valor máximo de la correlación se da cuando el ángulo que forman es $\alpha - \beta$, y el valor mínimo cuando el ángulo es $\alpha + \beta$. Los valores máximo y mínimo de la correlación son

$$\cos(\alpha - \beta) = \cos \alpha \cos \beta + \sin \alpha \sin \beta = \cos \alpha \cos \beta + \sqrt{(1 - \cos^2 \alpha)(1 - \cos^2 \beta)}$$

$$\cos(\alpha + \beta) = \cos \alpha \cos \beta - \sin \alpha \sin \beta = \cos \alpha \cos \beta - \sqrt{(1 - \cos^2 \alpha)(1 - \cos^2 \beta)}$$



y como $\cos\alpha=r_{12}$ y $\cos\beta=r_{13}$, tenemos $r_{23} \in r_{12}r_{13} \pm \sqrt{(1-r_{12}^2)(1-r_{13}^2)}$. Este es el caso particular de 6 cuando hay sólo tres variables.

8. CONCLUSIÓN

El método aquí presentado tiene la ventaja de que nos permite ir decidiendo los valores que queremos asignar a cada una de las correlaciones de la matriz que vamos a generar, de tal forma que conocemos en cada caso entre qué valores ha de encontrarse cada una de ellas para que la matriz final sea plausible. La plausibilidad viene dada por el hecho de que la correlación múltiple al cuadrado sea menor o igual a 1.

Por otra parte, en cada paso del proceso es preciso obtener la inversa de una matriz de correlaciones. Otra ventaja de este procedimiento consiste en que la inversa de la matriz de correlaciones de un orden dado se obtiene a partir de la inversa de la matriz de orden inmediatamente anterior. Eso simplifica enormemente los cálculos ya que se trata de un proceso recursivo fácilmente programable en un lenguaje de alto nivel.

9. REFERENCIAS

- BENTLER, P. M. (1989): *EQS, Structural Equation Program Manual*. BMDP Statistical Software. Los Ángeles, CA.
- HARRIS, R. J. (1985): *A primer of Multivariate Statistics (2nd Ed.)*. Academic Press, Inc. Orlando, Florida.

INTERPRETACIÓN DE LA FIABILIDAD EN EL ANÁLISIS DOCUMENTAL MEDIANTE ELECCIONES DICOTÓMICAS Y SU INCIDENCIA EN LA INVESTIGACIÓN CUALITATIVA

Santiago Nieto Martín
Universidad de Salamanca

En el ámbito de la investigación educativa (y, en general, de las ciencias sociales), cuando de planteamientos metodológicos se trata, nos encontramos, en demasiadas ocasiones, ante una serie de dilemas intuitivamente explicados/resueltos por parte del investigador, que, más que puramente técnicos, tienen su anclaje en cierta contaminación formativa del propio ambiente investigador, y que, en todo caso, «facilitan» el proceso de trabajo.

En ese sentido, resulta interesante considerar las ideas manifestadas por Cicourel (1982), de quien Habermas (1988: 182) se hace eco en los siguientes términos: ...«nuestra falta de refinamiento metodológico significa que los procedimientos de decisión en orden a categorizar los fenómenos sociales están llenos de supuestos implícitos de sentido común acerca del actor, las personas concretas, y de los propios puntos de vista del observador acerca de la vida diaria. Los procedimientos parecen intuitivamente «correctos» o «razonables» porque están enraizados en la vida diaria».

Tal expresión adquiere singular relevancia cuando se trata de metodologías aplicadas a las distintas manifestaciones del comportamiento humano. La aplicación de técnicas de análisis documental en la investigación educativa requiere refinamientos técnicos, insuficientemente resueltos, a los que pretendemos aportar nuestra reflexión teórico-práctica en torno a un aspecto clave como es la fiabilidad de los análisis que se realizan.

1. LA FIABILIDAD Y VALIDEZ EN EL ANÁLISIS DOCUMENTAL

Muchas decisiones de trascendencia en educación se basan en análisis informales que carecen de fiabilidad y validez a causa de un alto nivel de subjetividad y de la ausencia de un enfoque sistemático en el análisis. Aunque en la medición educativa, fiabilidad y validez están relacionadas entre sí por diversas proposiciones, en el presente trabajo pretendemos poner de manifiesto algunas cuestiones metodológicas para la interpretación más adecuada de lo que debe ser la fiabilidad referida al análisis documental aplicado a la investigación educativa.

«La importancia de la fiabilidad procede de la seguridad que ofrece en cuanto a que los datos han sido obtenidos con independencia del suceso, instrumento o persona que los mide. Por definición, los datos fiables son aquellos que permanecen constantes en todas las variaciones del proceso de medición» (Kaplan y Goldsen, 1965: 83). La fiabilidad, por tanto, mide el grado en el cual, cualquier diseño

de investigación (total o parcialmente) o los datos resultantes de la misma, «representan variaciones en los fenómenos reales, en lugar de representar las circunstancias extrínsecas de la medición, las idiosincrasias ocultas de cada uno de los analistas o las tendencias subrepticias de un procedimiento» (Krippendorff, 1990: 192).

Sin duda, tenemos que admitir que para verificar la fiabilidad del análisis documental *se requiere una considerable duplicidad de esfuerzos*. En dichos análisis distinguimos, al menos, tres tipos distintos de fiabilidad: estabilidad, reproducibilidad y exactitud. Cada uno de ellos requiere un diferente diseño, pero, en cuanto a su eficacia, podemos establecer tres niveles:

ESTABILIDAD < REPRODUCIBILIDAD < EXACTITUD

El diseño que requiere el análisis de la fiabilidad como *estabilidad*, parte del comúnmente denominado «test-retest», mediante el cual, no obstante, aceptamos el razonable supuesto de que se pueden poner de manifiesto las sucesivas incongruencias del observador; por ello, se considera que es el menos eficaz de todos.

La fiabilidad como *reproducibilidad* («test-test»), requiere, al menos, dos observadores. Ello no evita los errores de incongruencia que puedan cometer los observadores, además del desacuerdo que se puede producir entre ellos. No obstante, es más eficaz que el anterior, siendo, a partir de este nivel, donde plantearemos nuestra reflexión sobre el tema.

Sin duda, el más eficaz de los tipos de fiabilidad mencionados es el de *exactitud*. Para ello se establece un diseño «test-norma», donde sin embargo, no quedan eliminadas ni las incongruencias de los observadores, ni los desacuerdos entre ellos, y, donde, además, pueden aparecer desviaciones sistemáticas respecto a la norma.

2. ÍNDICE DE FIABILIDAD: CÁLCULO E INTERPRETACIÓN

Los datos sobre fiabilidad exigen que dos codificadores, como mínimo, describan de forma independiente un conjunto, probablemente elevado, de unidades de registro en términos de un lenguaje de datos común, que bien pueden ser en términos de categorías descriptivas, valores numéricos de una variable o, en muchas ocasiones, complejos esquemas clasificatorios.

Podemos considerar que la fiabilidad está garantizada cuando el acuerdo entre los codificadores es total; sin embargo, *si el acuerdo no es mayor que el correspondiente al azar, la fiabilidad es nula*. El problema radica pues en «medir» adecuadamente el acuerdo alcanzado por los observadores, codificadores o jueces al procesar de forma independiente la información científica que se somete a estudio e investigación.

Aunque nosotros vamos a exponer sobre un ejemplo concreto una metodología adecuada para valorar correctamente la fiabilidad de un análisis, hemos de manifestar que el hecho de que dos observadores (o más) *coincidan en un determinado porcentaje de acuerdo* en la asignación de los datos a las categorías, *no tiene ningún significado especial*. Y esto es muy importante considerarlo por cuanto en numerosos manuales sobre análisis documental se utiliza, como criterio de fiabilidad, el coeficiente porcentual alcanzado, y, sin duda, así considerado, resulta engañoso.

Supongamos, a modo de ejemplo, que queremos conocer ciertas características psicológicas de una persona a través del análisis de algunos documentos escritos y pertenecientes a esa persona. Para ello, dos codificadores con la debida preparación y formación para realizar dicho análisis, reciben el encargo de tratar de identificar una característica determinada en 15 documentos, marcándolos con el dígito «0» en el supuesto de no identificar nada relacionado con la instrucción recibida, o bien, con un «1», en caso contrario.

Una vez realizado el mencionado trabajo, los datos de la situación «test-test» generada, se pueden representar de la siguiente manera:

Documentos	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Codificador A	0	0	1	1	0	0	0	1	0	1	1	1	0	0	1
Codificador B	0	0	0	1	0	1	0	1	0	1	0	1	0	0	0

El simple análisis de los datos muestra que los dos codificadores (u observadores) coinciden en once de un total de quince asignaciones (es decir, en el 73,3 %). Ahora bien, ese porcentaje que, en principio, no tiene por qué «explicarnos» mucho, *nada nos dice sobre si esa cifra es grande o pequeña, y, sobre todo, de que manera puede compararse con lo que hubiera sucedido por obra del azar*. Posiblemente no podamos dar respuesta convincente a esos interrogantes, pero, en cualquier caso, siempre hemos de tenerlos en cuenta. Para ello, los datos obtenidos en la matriz 2 x 15, los colocamos en una tabla de contingencia de frecuencias de 2 x 2. Así:

		Codificador B		
		0	1	
Codificador A	0	7	1	8
	1	3	4	7
		10	5	15

Deseamos obtener, a partir de los datos anteriores, una tabla semejante en la que el acuerdo sea producto del azar. Aunque ignoramos esa distribución, podemos estimarla a partir de la «experiencia» mostrada por los dos codificadores.

El Codificador A sostiene que ha identificado siete documentos conforme a la instrucción recibida, de un total de quince; sin embargo, el Codificador B mantiene que identificó cinco del total de quince. Por lo tanto, en conjunto, podemos decir que identificaron doce de un total de treinta (o sea, el 40 %: ese porcentaje se corresponde exactamente con el número de «1» en la matriz de 2 x 15).

Ahora bien, si los dieciocho «0» y los doce «1» los introdujéramos en una urna y fueran extraídos al azar por los dos codificadores, nos encontraríamos con las siguientes situaciones:

— La probabilidad de extraer dos «0» sin reposición (uno por cada observador), sería la siguiente: $(18/30) (17/29) = 0,351$; multiplicada esa cantidad por el número de unidades de registro (en nuestro caso, 15), obtenemos un valor aproximado de 5,3.

— Del mismo modo, al extraer dos «1», ésta sería la situación: $(12/30) (11/29) = 0,151$, que, multiplicado por 15, nos daría un valor aproximado de 2,3.

— Así mismo ocurriría al extraer «0» y «1»: $(18/30) (12/29) = 0,248$, con un valor aproximado de 3,7 al multiplicarlo por 15. Igual valor obtendríamos al extraer «1» y «0», respectivamente.

Es evidente que si se produjera un acuerdo perfecto, sólo «encontraríamos» valores en la diagonal principal, es decir, la que va del casillero 0-0 al casillero 1-1, dado que todo lo demás sería nulo; en ese sentido, el número máximo de veces que, sin reposición, obtendríamos el par 0-0, sería 9; así mismo, el número máximo de veces del par 1-1, sería 6. Por lo tanto, las tres tablas que nos han de servir de referencia para establecer una *medida adecuada de acuerdo* son:

9
6

Acuerdo máximo

7	1
3	4

Co-ocurrencias observadas

5,3	3,7
3,7	2,3

Acuerdo por azar

Se considera que la medida más convincente para expresar la cantidad de acuerdo en los datos sobre fiabilidad resulta del *grado de semejanza entre la tabla de co-ocurrencias observadas y la tabla de acuerdo máximo*, en lugar del grado en que el acuerdo es meramente producto del azar.

La fórmula mediante la que calculamos el coeficiente de acuerdo buscado (α) es la siguiente:

$$\text{Co-ocurrencias observadas} = \alpha \cdot \left(\begin{matrix} \text{acuerdo} \\ \text{máximo} \end{matrix} \right) + (1 - \alpha) \cdot \left(\begin{matrix} \text{acuerdo} \\ \text{por azar} \end{matrix} \right)$$

Cuando la tabla de co-ocurrencias observadas sea igual a la tabla de acuerdo máximo, entonces $\alpha = 1$; por el contrario, $\alpha = 0$ cuando la tabla de co-ocurrencias observadas sea igual a la tabla de acuerdo por azar. Por lo tanto, tras diversas operaciones algebraicas, llegamos a la expresión:

$$\alpha = 1 - \frac{\text{discrepancia observada}}{\text{discrepancia prevista}}$$

Dado que en la tabla 2 x 2 las discrepancias aparecen en los casilleros 0-1 y 1-0, podemos establecer:

- Discrepancia observada: $3 + 1 = 4$
- Discrepancia prevista: $3,7 + 3,7 = 7,4$; luego:

$$\alpha = 1 - \frac{(3 + 1)}{(3,7 + 3,7)} = 0,4594$$

El coeficiente obtenido nos lleva a deducir que el acuerdo alcanzado resulta ser un 45,94 % por encima del obtenido al azar, lo que pone de manifiesto que los dos codificadores han leído e identificado correctamente la referencia señalada sólo el 45,95 % de todos los documentos. Este dato pone de relieve la necesidad de alejarse del uso indiscriminado de los acuerdos porcentuales como criterio de fiabilidad.

El coeficiente de acuerdo indica de qué manera es posible reconstruir la tabla original a partir de las dos tablas de acuerdo máximo y de acuerdo por azar, respectivamente:

A partir del acuerdo máximo: 0,4594

9
6

=

4,1
2,8

A partir del acuerdo por azar: (1 - 0,4594)

5,3	3,7
3,7	2,3

=

2,9	2
2	1,2

Tabla simétrica equivalente a la observada:

7	2
2	4

Las tablas siguientes son todas equivalentes entre sí en cuanto al grado de acuerdo/discrepancia:

7 4	7 3	7 2	7 1	7 0
0 4	1 4	2 4	3 4	4 4

Para finalizar, hemos de hacer necesaria referencia a que la cuestión planteada es la más elemental de cuantas pueden ofrecerse (dos codificadores y datos dicotómicos). Sin duda, resulta evidente pensar que la cuestión se complica a partir de lo que podemos considerar más normal en cualquier proceso de investigación; es decir, dos o más codificadores, y, sobre todo, si se amplía el número de categorías, o, incluso, el tipo de métrica. De ahí la prudencia que debemos conceder a la interpretación de coeficientes que en cada momento podamos obtener.

REFERENCIAS BIBLIOGRÁFICAS

- BARDIN, L. (1986): *Análisis de Contenido*. Madrid: Akal Universitaria.
 BERELSON, B. (1952): *Content analysis in communication research*. Glencoe: The Free Press.
 CICUOREL, A. (1982): *El método y la medida en sociología*. Madrid: Editora Nacional.
 FOX, D. (1981): *El proceso de investigación en educación*. Pamplona: EUNSA.
 HABERMAS, J. (1988): *La lógica de las ciencias sociales*. Madrid: Ed. Tecnos.
 KRIPPENDORFF, K. (1990): *Metodología de análisis de contenido*. Barcelona: Ed. Paidós Ibérica.

LA UTILIZACIÓN DE LOS PROCEDIMIENTOS DE COMPARACIONES MÚLTIPLES EN LA INVESTIGACIÓN EDUCATIVA EN ESPAÑA

A. Sáez, J. M. Suárez, F. Aliaga y R. M. Bo
Universitat de València¹

1. PRESENTACIÓN DEL TRABAJO

Un problema común al que nos podemos enfrentar en cualquier investigación es querer comparar más de 2 grupos de datos para detectar posibles diferencias entre ellos. La utilización de modelos de ANOVA puede permitirnos detectar diferencias, a nivel global, entre las medias involucradas, pero en muchas ocasiones deseamos trabajar a un mayor detalle y detectar las diferencias entre grupos concretos lo que sólo es posible mediante el uso de los Procedimientos de Comparaciones múltiples (PCM).

Las propiedades estadísticas de los PCM son bien conocidas (Miller, 1981; Hochberg y Tamhane, 1987) y el comportamiento de estas pruebas en distintas situaciones ha sido descrita tanto en estudios teóricos (Tukey, 1953; Einot y Gabriel, 1975; Stoline y Ury, 1979; Scheffé, 1970) como en estudios utilizando estrategias de simulación de Monte-Carlo (Dunnnett, 1980a; Dunnnett, 1980b; Wang, 1971; Maxwell, 1980; Keselman y Rogan, 1978; Keselman, Toothaker y Shooter, 1975; Keselman, Murray y Rogan 1976; Berhardson, 1975; Boardman y Moffitt, 1971). A pesar de todo esto, las Comparaciones Múltiples no suelen ser utilizadas por los investigadores o, en algunas ocasiones se utilizan incorrectamente.

La posibilidad de efectuar comparaciones múltiples ha recibido críticas importantes desde diversas perspectivas (Dawkins, 1983; O'Neill, y Wetherill, 1971; Perry, 1986). Bastantes críticas provienen de la pugna teórica entre las diversas concepciones y corrientes estadísticas. Otras se refieren a principios de utilización incorrecta. Por ejemplo, Wilcox (1987) afirma que solamente cuando se desean realizar todas las comparaciones por pares y si se quiere mantener la probabilidad del error de tipo I igual a α , es perfectamente legítimo omitir la prueba F y usar uno de los procedimientos de contraste de medias. La razón aducida es que la aplicación de las pruebas de comparación múltiple únicamente después de una prueba F significativa es una estrategia que reduce la potencia y el nivel α en una cantidad difícil de determinar.

* Avda. Blasco Ibáñez, 21, 46010-VALENCIA. Teléfono 3864430. Centralita 3864420, ext. 6245.

¹ Por lo que se refiere a manuales, solamente se cuenta con algunos, el trabajo primigenio de síntesis de Miller, ya actualizado (Miller, 1981), y las recientes exposiciones en los textos de Klockars y Sax (1986) y Toothaker (1993) en la colección de Sage y el más completo de Hochberg y Tamhane (1987); ninguno de ellos traducido al castellano.

La utilización incorrecta de Comparaciones Múltiples lleva a los investigadores a conclusiones erróneas que se reflejan en ambos tipos de error (tipo I y tipo II). En un trabajo de Coward (1991) sobre la utilización de las pruebas de comparaciones múltiples en Estados Unidos se detectan cuatro posibles situaciones que pueden conducir a error en la aplicación de las pruebas: 1) utilizar pruebas de comparaciones de pares cuando lo correcto es utilizar contrastes polinómicos, 2) usar comparaciones múltiples a posteriori en lugar de a priori; 3) utilizar medias aritméticas en lugar de mínimo cuadráticas y 4) utilizar una prueba demasiado «liberal».

A estos problemas debidos a la mala utilización, hay que añadir la falta de uso de este tipo de pruebas que, como veremos, se da en nuestro país y que a nuestro juicio se da por dos circunstancias:

1) La falta de claridad en los textos de estadística sobre los distintos procedimientos de comparaciones múltiples con una clara clasificación de las pruebas en sus aspectos más relevantes como: conveniencia respecto al diseño experimental utilizado, tratamiento del control del error de tipo I, o incluso a nivel de los supuestos estadísticos necesarios para su aplicación. Y como consecuencia de lo anterior,

2) La falta de implementación en paquetes estadísticos de ordenador de muchas de las pruebas para los diseños experimentales más utilizados. Aunque todos los paquetes suelen cubrir los diseños de una vía entre grupos, hemos detectado una carencia de pruebas de comparaciones múltiples para diseños factoriales de medidas repetidas o en los diseños mixtos o en los diseños de ANOVA no paramétricos.

Existen diferentes tipos de clasificaciones de las pruebas de comparaciones múltiples lo cual nos ofrece una variedad de dimensiones según las cuales caracterizar a las pruebas (Hochberg y Tamhane, 1987; Toothaker, 1991). Por ejemplo, Toothaker (1991, 1993) propone las siguientes dimensiones: 1) **Según el número de comparaciones**; 2) **según el tipo de contraste** (ortogonales frente a no ortogonales); 3) **según la manera de comparación** (por pares o no); 4) **comparaciones a priori o comparaciones a posteriori**; 5) **según el proceso de cálculo** (simples o en un único paso o en varios pasos 'stepwise', estos a su vez se dividen en *step-down* o *step-up*, según se proceda desde la mayor diferencia hasta la más pequeña o desde la menor diferencia a la mayor); 6) **según el tipo de estadístico y/o la distribución teórica utilizada en su cálculo**; 7) **según el tipo de tasa de error** (existen dos tipos: tasa de error por comparación y tasa de error por familia).

Muchas de las dimensiones anteriores pueden aparecer combinadas según la elección de la prueba que se realice. Se podrían utilizar comparaciones múltiples a priori y una tasa de error por comparación. O se pueden utilizar comparaciones ortogonales y a posteriori con una tasa de error por comparación. Algunas combinaciones son imposibles, tales como todas las comparaciones por pares y ortogonales. Sin embargo, es cierto que unas determinadas combinaciones se suelen utilizar con más frecuencia, como comparaciones ortogonales y a priori con una tasa de error por comparación.

A modo de cuadro-resumen presentamos algunas de las pruebas de comparaciones múltiples clasificadas según la distribución estadística que utilizan en su cálculo:

Basadas en la distribución t	Dunn-Bonferroni (Dunn, 1961) Dunn-Sidak (Dunn, 1958 y Sidák, 1967) Holm-Shaffer (Holm, 1979 y Shaffer, 1986)
Basadas en la distribución del Rango Studentizado	Tukey (Tukey, 1953) Newman-Keuls (Newman, 1932 y Keuls 1952) Duncan (Duncan, 1955) Ryan (Ryan, 1960; Einot y Gabriel , 1975) Peritz (Peritz, 1970)
Basadas en la distribución F	Scheffé (Scheffé, 1953, 1959) F de Newman-Keuls F de Ryan
Basadas en una prueba t protegida	LSD de Fisher (Fisher, 1935) Shaffer-Ryan (Shaffer, 1979) Fisher-Hayter (Hayter, 1986)
Basadas en la comparación con un control	Dunnet (Dunnett, 1955)

En el resumen anterior podemos observar que, además de las pruebas tradicionales, existen otros procedimientos más recientes (como los de Ryan o Peritz) que suelen ser modificaciones de pruebas anteriores para corregir algunos de los problemas de estas pruebas clásicas. Se da la circunstancia que muchas de estas pruebas recientes no están disponibles en los manuales de estadística al uso y, por supuesto, tampoco están implementadas en los paquetes estadísticos para ordenador más utilizados.

Como ocurre en la prueba t y la prueba F es necesario el cumplimiento de los supuestos paramétricos para la correcta aplicación de las pruebas de contraste que hemos visto. Vamos a revisar aquí algunos de los resultados más importantes que se han encontrado sobre el cumplimiento de supuestos.

Respecto al supuesto de normalidad parece que según Dunnet (1982), y como ocurre con la prueba F, las pruebas de contraste son robustas frente a pequeñas desviaciones respecto a la normalidad; en el caso de desviaciones de la normalidad muy grandes, aumenta el riesgo del error (consultar a Dunnet, 1982 y Ringland, 1983).

Respecto al problema de tamaños muestrales distintos entre los grupos a comparar se proponen distintas pruebas alternativas como son dos variaciones a la prueba de Tukey: una propuesta por Kramer (1956), conocida como prueba de Tukey-Kramer, y la segunda variación propuesta por Miller (1981) y Winer (1971), conocida como prueba de Miller-Winer y por último una prueba nueva propuesta por Hochberg (1974) conocida como la prueba GT2.

Mientras que para el caso de desigualdad de varianzas existen alternativas como la prueba GH de Games y Howell (1976), y las pruebas C o T3 de Dunnet (1980). Brown y Forsythe (1974c) han propuesto una modificación a la prueba de Scheffé para hacerla resistente a la desigualdad de las varianzas y que ha dado pie a posteriores variaciones (Kaiser y Bowden, 1983) y alternativas (Dalal, 1975; Hochberg, 1976).

En este trabajo se pretende llevar a cabo un contraste entre estas posibilidades técnicas y la realidad de utilización concreta en nuestro ámbito de investigación. Así, pretendemos determinar en que situaciones se emplean estos procedimientos, en cuales se podrían/deberían emplear, qué opciones

concretas se manejan, etc. Todo ello vamos a llevarlo a cabo dentro del ámbito de la investigación educativa en nuestro país.

2. MÉTODO

A tal efecto se ha tomado como referente de investigación los trabajos publicados en revistas de investigación de difusión nacional. En este sentido, se han seleccionado 7 publicaciones: Revista de Investigación Educativa, Infancia y Aprendizaje, Revista de Educación, Investigación en la Escuela, Bordón, Revista Española de Pedagogía y Ciencias de la Educación. Dada la relativa recencia de estos procedimientos como tema monográfico en la literatura estadística se ha seleccionado un conjunto de 5 años correspondientes al período que va desde 1988 a 1992, ambos inclusive.

La técnica de trabajo es la correspondiente a cualquier estudio bibliométrico básico, procurando mantener en todo momento la conexión con las informaciones cualitativas que se derivan de los informes originales.

3. RESULTADOS

El conjunto de resultados respecto a la utilización de las pruebas de Comparaciones Múltiples se encuentran recogidos en las tablas 1 y 2. Es preciso resaltar que se indagaron únicamente tres publicaciones (Revista de Investigación Educativa, Infancia y Aprendizaje y Bordón) por ser las únicas de las revistas estudiadas en las que se encuentran artículos que utilizan las pruebas de Comparaciones Múltiples.

A partir del análisis de esta información se pueden señalar los siguientes aspectos relevantes:

- Existe esencialmente una revista —Infancia y Aprendizaje— en la que es relativamente habitual la publicación de trabajos que incluyen la utilización de estas pruebas. Y aún así su presencia es relativamente moderada respecto a las posibilidades potenciales directas para su empleo: un 38,89% de los artículos que utilizan modelos ANOVA. En las otras dos revistas la publicación de trabajos que presenten Comparaciones Múltiples es mucho más rara.
- No parece existir ningún tipo de evolución temporal a través de los cinco años estudiados respecto a la utilización de las pruebas de Comparación Múltiple. De hecho, los escasos trabajos que las utilizan tienden a distribuirse de una forma casi uniforme a través de todo este período temporal.
- Merece un comentario detallado la escasa utilización de estas pruebas incluso tomando como referente las situaciones en las que es perfectamente ajustada su utilización: aquellos trabajos en los que se ha utilizado modelos de ANOVA. Así, sólo un 13,64% de los trabajos de las tres revistas, entre los que emplean modelos ANOVA, ha utilizado alguna técnica de Comparaciones Múltiples. Además, sí que parece haber una cierta relación entre la mayor utilización de estas técnicas y su mayor actuación proporcional respecto al total de situaciones posibles. Así, en ambos casos la revista Infancia y Aprendizaje muestra tanto una mayor frecuencia de utilización como una mayor proporción de utilización respecto a todos los trabajos que emplean modelos ANOVA.

TABLA 1

NÚMERO DE ARTÍCULOS CON APLICACIONES DE PRUEBAS DE COMPARACIONES MÚLTIPLES PARA CADA REVISTA Y AÑO. SE ESTABLECEN PROPORCIONES RESPECTO A LA TOTALIDAD DE LOS ARTÍCULOS PUBLICADOS, RESPECTO A LA TOTALIDAD DE ARTÍCULOS DE LA REVISTA Y RESPECTO A LA TOTALIDAD DE ARTÍCULOS POR AÑOS —EN LOS TRES CASOS ES LA TOTALIDAD DE ARTÍCULOS QUE UTILIZAN ESTAS PRUEBAS—. (T= TUKEY, S=SCHEFFÉ Y B= BONFERRONI)

1988 1989	1990	1991	1992	Total		
Revista Investigación Educativa	0	0	0	1(T)	0	1
% respecto Total	0,0	0,0	0,0	11,11%	0,0	11,11
% respecto Revista	0,0	0,0	0,0	100%	0,0	
% respecto Año	0,0	0,0	0,0	33,33%	0,0	
Infancia y Aprendizaje	2(T)	1(S)	1(B)	2(S)	1(S)	7
% respecto Total	22,22	11,11	11,11	22,22	11,11	77,78
% respecto Revista	28,57	14,29	14,29	28,57	14,29	
% respecto Año	100	100	50	66,67	100	
Bordón	0	0	1(S)	0	0	1
% respecto Total	0,0	0,0	11,11	0,0	0,0	11,11
% respecto Revista	0,0	0,0	100	0,0	0,0	
% respecto Año	0,0	0,0	50	0,0	0,0	
Total	2	1	2	3	1	9
% Total	22,22	11,11	22,22	33,33	11,11	

- Los trabajos que emplean Comparaciones Múltiples tienden a utilizar aquellas que se corresponden con los modelos más simples —ANOVAS de una vía— con técnicas recogidas en los paquetes estadísticos más extendidos (SPSS, BMDP, fundamentalmente). Una parte de las «exclusiones» en cuanto a la utilización de estos procedimientos se produce por la presencia de modelos ANOVA más complejos —con 2 ó más variables independientes—. Esto parece tener una clara relación con la mucho menor oferta de estos procedimientos en los paquetes estadísticos, junto con un tratamiento más esporádico del problemas en los textos y manuales dedicados a esta temática. Así, salvo los manuales que tratan monográficamente el tema de las técnicas de Comparaciones Múltiples, las opciones y forma de aplicación de las mismas a los modelos de cierta complejidad no están tradicionalmente recogidas en los textos, salvo alguna mención tangencial en algunos casos.

TABLA 2

NÚMERO DE ARTÍCULOS CON APLICACIONES DE PRUEBAS DE COMPARACIONES MÚLTIPLES PARA CADA REVISTA, PROPORCIONES SOBRE EL TOTAL DE ARTÍCULOS QUE HAN EMPLEADO UN MODELO ANOVA Y SOBRE EL TOTAL DE ESTUDIOS EN QUE SE UTILIZA EL ANÁLISIS CUANTITATIVO

	1988	1989	1990	1991	1992	Total
Revista Investigación Educativa	0	0	0	1(T)	0	1
% respecto Trabajos estadísticos	0,0	0,0	0,0	11,11%	0,0	1,33
% respecto ANOVAS	0,0	0,0	0,0	100	0,0	8,33
Infancia y Aprendizaje	2(T)	1(S)	1(B)	2(S)	1(S)	7
% respecto Trabajos estadísticos	14,29	5,88	8,33	14,29	11,11	10,61
% respecto ANOVAS	100	12,5	25	66,67	100	38,89
Bordón	0	0	1(S)	0	0	1
% respecto Trabajos estadísticos	0,0	0,0	16,67	0,0	0,0	1,72
% respecto ANOVAS	0,0	0,0	100	0,0	0,0	9,09
Total	2	1	2	3	1	9
% respecto Trabajos estadísticos	3,45	1,12	1,9	5	2	2,49
% respecto ANOVAS	22,22	6,67	8,7	30	11,11	13,64

- Del total de 9 artículos en los que se emplean técnicas de Comparaciones Múltiples más de la mitad (55,56%) utilizan la prueba de Scheffé. En los casos restantes, se utilizan en una tercera parte (3 artículos) la prueba de Tukey (%) y en un sólo artículo (%) la prueba de Bonferroni. A partir de las informaciones sobre los estudios respecto a las propiedades de estas pruebas las decisiones no parecen estar mayoritariamente respaldadas por la evidencia. Toothaker (1993) recomienda la utilización de la prueba de Tukey por no resultar tan conservadora como la de Scheffé ni tan liberal como la de Bonferroni. Por su parte, HOCHBERG Y TAMHANE (1987) señalan que la prueba más potente con diseños equilibrados es la de Scheffé, mientras que la de Tukey es más adecuada cuanto más se acentúa el desequilibrio entre los grupos. Dado que la práctica totalidad de los estudios manejan grupos desequilibrados parece que se ha producido mayoritariamente una decisión no suficientemente avalada, respecto a la técnica concreta elegida para establecer las Comparaciones Múltiples, en dos terceras partes de los trabajos estudiados (66,66%).
- Por lo que se refiere a la utilización de las pruebas dentro del enfoque del ajuste de modelos, es preciso resaltar que se produce en la práctica totalidad de las situaciones una aplicación independiente de la verificación del cumplimiento de los supuestos del modelo. Este hecho, por otra parte, es algo desgraciadamente habitual en el caso de la aplicación de múltiples técnicas estadísticas y particularmente se tiende a producir en los casos en que se utilizan modelos ANOVA.

4. CONCLUSIONES

La revisión de la literatura científica más reciente respecto a las técnicas estadísticas está acentuando cada vez más la importancia de la utilización de los procedimientos de Comparaciones Múltiples cuando se trata de verificar hipótesis sobre la igualdad de K medias correspondientes a situaciones diferentes que se desean comparar.

Existe hoy en día suficientes alternativas desarrolladas dentro de las técnicas de Comparaciones Múltiples que pueden dar respuesta a buena parte de las necesidades más comunes en la investigación educativa. De hecho, se cubre la mayor parte de los modelos que se aplican a las situaciones que habitualmente se plantean en nuestro ámbito de investigación.

No obstante, a partir de la revisión de los trabajos de investigación educativa en el ámbito español se aprecia que, en conjunto, el tema de la aplicabilidad de estas técnicas está poco difundido en la comunidad científica. De hecho, se manejan los procedimientos más clásicos que son de amplia difusión en los textos de estadística y que se encuentran profusamente recogidos en los paquetes estadísticos más importantes. Además, se emplean exclusivamente aquellas técnicas relacionadas con los modelos más simples, con una sola variable independiente y en situaciones entre grupos. Por último, cabe señalar que las decisiones entre las opciones alternativas no se suelen hacer tomando como base los estudios de validación disponibles respecto a los diferentes procedimientos.

Por todo ello, es preciso afirmar la necesidad de establecer los mecanismos para la difusión de este procedimiento entre la comunidad científica que investiga en educación. Esto entendemos que debe hacerse realizando un esfuerzo por clarificar las opciones disponibles, establecer estrategias de adecuación de estas opciones a cada situación concreta y reseñar cuales son las herramientas informáticas disponibles en cada situación.

Finalmente, pensamos que sería de particular interés en este tema la elaboración de programas informáticos específicos que cubrieran las principales lagunas existentes y que permitieran una utilización más amplia y mejor dirigida de estos procedimientos, a la espera de su implementación en los paquetes estadísticos más conocidos.

5. BIBLIOGRAFÍA

- BERHARDSON, C. (1975): Type I Error Rates when Multiples Comparison Procedures Follow a Significant F Test of ANOVA. *Biometrics*, **31**, 229-232.
- BOARDMAN, T. & MOFFITT, D. (1971): Graphical Monte Carlo Type I Error Rates for Multiple Comparison Procedures. *Biometrics*, **27**, 728-744.
- BROWN, M. & FORSYTHE, A. (1974c): The ANOVA and Multiple Comparisons for Data with Heterogeneous Variances. *Biometrics*, **30**, 179-184.
- COWARD, W. M. (1991): *A Meta-Analysis of Multiple Comparison Procedures*. Tesis doctoral.
- DALAL, S. (1975): Simultaneous Confidence Procedure for Univariate Behrens-Fisher Type Problems. *Biometrics*, **65**, 221-225.
- DAWKINS, H. (1983): Multiple Comparisons Misused: Why so Frequently in Response-Curve Studies? *Biometrics*, **39**, 789-790.
- DUNCAN, D. (1955): Multiple Range Tests and Multiple F Test. *Biometrics*, **11**(1), 1-42.
- DUNN, O. J. (1958): Estimation of the Means of dependent variables. *Annals of Mathematical Statistics*, **29**, 1.095-1.111.
- DUNN, O. J. (1961): Multiple comparisons using rank sums. *Technometrics*, **6**, 241-252.
- DUNNETT, C. (1955): A Multiple Comparison Procedure for Comparing Several Treatments with a Control. *Journal of the American Statistical Association*, **50**, 1.096-1.121.

- DUNNETT, C. (1980a): Pairwise Multiple Comparisons in the Homogeneous Variance, Unequal Sample Size Case. *Journal of the American Statistical Association*, **75**(372), 789-795.
- DUNNETT, C. (1980b): Pairwise Multiple Comparisons in the Unequal Variance Case. *Journal of the American Statistical Association*, **75** (372), 796-800.
- DUNNETT, C. (1982): Robust multiple comparisons. *Communications in Statistics*, **11**, 2.611-2.629.
- EINOT, I. & GABRIEL, K. (1975): A Study of the Powers of Several Methods of Multiple comparisons. *Journal of the American Statistical Association*, **70** (351), 574-583.
- FISHER, R. A. (1935): *The design of experiments*. Edinburgh: Oliver & Boyd.
- GAMES, P. A. & HOWELL, J. F. (1976): Pairwise Multiple comparisons procedures with unequal n's and/or variances. *Journal of Educational Statistics*, **1**, 113-125.
- HAYTER, A. J. (1986): The maximum familywise error rate of Fisher's least significant difference test. *Journal of the American Statistical Association*, **81**, 1.000-1.004.
- HOCHBERG, Y. (1974): Some generalizations of the T-method in simultaneous inference. *Journal of Multivariate Analysis*, **4**, 224-234.
- HOCHBERG, Y. & TAMHANE, A. (1987): *Multiple Comparison Procedures*. New York: John Wiley & Sons.
- HOCHBERG, Y. (1976): A Modification of the T-method of Multiple comparisons for a One-Way Layout with Unequal Variances. *Journal of the American Statistical Association*, **71** (353), 200-203.
- HOLM, S. (1979): A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, **6**, 65-70.
- KAISER, L. & BOWDEN, D. (1983): Simultaneous Confidence Intervals for all Linear Contrast of Means with Heterogeneous Variances. *Communications in Statistics-Theory and Methods*, **12**, 73-88.
- KESELMAN, H. & ROGAN, J. (1978): A Comparisons of the Modified-Tukey and Scheffé Methods of Multiple comparisons for pairwise Contrasts. *Journal of the American Statistical Association*, **73** (361), 47-52.
- KESELMAN, H., MURRAY, R. & ROGAN, J. (1976): Effect of Very Unequal Group Sizes on Tukey's Multiple Comparison Test. *Educational and Psychological Measurement*, **36**, 263-270.
- KESELMAN, H., TOOTHAKER, L. & SHOOTER, M. (1975): An Evaluation of Two Unequal n(k) Forms of the Tukey Multiple Comparison Statistic. *Journal of American Statistical Association*, **70**(351), 584-587.
- KEULS, M. (1952): Use of the «Studentized Range» in Connection with an Analysis of Variance. *Euphytica*, **1**, 112-122.
- KLOCKARS, A. J. & SAX, G. (1986): *Multiple Comparisons*. Beverly Hills: Sage.
- KRAMER, C. (1956): Extension of Multiple Range Tests to Group Means with Unequal Numbers of replications. *Biometrics*, **12**, 307-310.
- MAXWELL, S. (1980): Pairwise Multiple Comparisons in Repeated Measures Designs. *Journal of Educational Statistics*, **5** (3), 269-287.
- MILLER, R. (1981): *Simultaneous Statistical Inference* (2 ed.). New York: Springer-Verlag.
- NEWMAN, D. (1939): The Distribution of Range in Samples from a Normal Population, Expressed in Terms of an Independent Estimate of Standard Deviation. *Biometrika*, **31**, 20-30.
- O'NEILL, R. T. & WETHERILL, B. G. (1971): The present state of multiple comparisons methods (with discussion). *Journal of the Royal Statistical Society*, **33**, 218-241.
- PERITZ, E. (1970): *A note on multiple comparisons*. Unpublished manuscript, Hebrew University, Israel.
- PERRY, J. (1986): Multiple Comparison Procedures: A Dissenting View. *Journal of Economic Entomology*, **79**(5), 1.149-1.155.

- RINGLAND, J. (1983): Robust Multiple Comparisons. *Journal of American Statistical Association*, **78**, 145-151.
- RYAN, T. A. (1960): Significance tests for multiple comparisons of proportions, variance and other statistics. *Psychological Bulletin*, **57**, 318-328.
- SCHEFFÉ, H. (1953): A method for judging all contrasts in analysis of variance. *Biometrika*, **40**, 87-104.
- SCHEFFÉ, H. (1959): *The analysis of variance*. New-york: John Wiley.
- SCHEFFÉ, H. (1970): Practical Solutions of the Behrens-Fisher Problem. *Journal of the American Statistical Association*, **65** (332), 1501-1508.
- SHAFFER, J. P. (1979): Comparison of means: An F test followed by a modified multiple range procedure. *Journal of Educational Statistics*, **4**, 14-23.
- SHAFFER, J. P. (1986): Modified sequentially rejective multiple test procedures. *Journal of the American Statistical Association*, **81**, 826-831.
- SIDÁK, Z. (1967): Rectangular confidence regions for the means of multivariate normal distributions. *Journal of the American Statistical Association*, **62**, 626-633.
- STOLINE, M. & URY, H. (1979): Tables of the Studentized maximum Modulus Distribution and an Application to Multiple Comparison among Means. *Technometrics*, **21**(1), 87-93.
- TOOTHAKER, L. E. (1991): *Multiple Comparisons for Researchers*. Newbury Park: Sage Pubns.
- TOOTHAKER, L. E. (1993): *Multiple Comparisons procedures*. Beverly Hills: Sage Pubns.
- TUKEY, J. (1953): *The Problem of Multiple Comparisons*. Unpublished manuscript. Princeton University.
- WANG, Y. (1971): Probabilities of the Type I Errors of the Welch Tests for the Behrens-Fisher Problem. *Journal of the American Statistical Association*, **66** (335), 605-608.
- WILCOX, R. R. (1987): *New statistical procedures for the social sciences*. Hillsdale, N.J.: LEA.
- WINER, B. J. (1971): *Statistical principles in experimental design (2nd ed.)*. New-York: McGraw-Hill.

CONSTRUCCIÓN DE PRUEBAS CRITERIALES DE AULA

Leonor Buendía Eisman, Honorio Salmerón Pérez
Universidad de Granada

I. MARCO TEÓRICO Y LEGAL

Partimos del supuesto de que con programas educativos adaptados a las peculiaridades de los alumnos se puede conseguir que la mayor parte de ellos alcancen los objetivos de un curso o materia de enseñanza.

A partir de este supuesto, las pruebas referidas a normas resultarían incongruentes para la efectividad de la instrucción ya que si la mediación del profesor con su programa funciona, la dispersión de las puntuaciones en las pruebas se reduce dificultando así la interpretación con referencia al grupo (de la Orden, 1982).

— A modo de recuerdo, Hambleton y Swaminathan en 1985 nos informaban de las posibles utilizaciones de las Evaluaciones Referidas a Criterio en el aula y que resumimos en breves líneas:

- Guiar el progreso individual en programas basados en objetivos de aprendizaje.
- Diagnosticar deficiencias de aprendizaje.
- Evaluar programas de acción educativa.
- Orientar competencias en diferentes exámenes y certificaciones.

—

Es en el primer uso, la evaluación de alumnos, donde enmarcamos este trabajo de construcción de pruebas criteriosales, como herramientas al servicio del propio proyecto educativo y del currículum en que se concreta.

De esta forma, creemos que estamos en consonancia con el modelo evaluativo de la LOGSE para este ámbito, (Pérez Juste, 1992, p. 42) ya que, pueden servir a una función formativa de la evaluación, ayudar a determinar el logro de objetivos y al final de cada ciclo utilizarlos para obtener información válida que sirva para decisiones de promoción y/o certificación.

No es nuestro cometido hacer una reflexión teórica sobre lo que son Tests de Referencia Criterial (TRC), que puede verse en la obra de Gray (1978) quien ya aportaba un trabajo sobre 57 definiciones; ni tampoco hablar de los precursores de este término, (Glaser, 1963; Popham y Husek, 1969), sino aportar una experiencia práctica llevada a cabo con profesores de aula de E. Primaria sin que por ello desmerezca el soporte metodológico que avala la calidad técnica para el uso que van a tener.

— Utilizamos el esquema de desarrollo propuesto por Hambleton, R. 1990, hasta la validación de contenido de las pruebas acordando con Beggs y Lewis, 1979, que dado el propósito de las mismas, al

realizar evaluaciones internas en clase con función formativa, lo más importante es demostrar su validez de contenido.

II. DESARROLLO

1. Consideraciones preliminares a la construcción de las pruebas

a) Declaración de los propósitos de las pruebas. Es nuestra intención, dentro de una línea de investigación de nuestro Departamento, fomentar la colaboración de los centros de nuestro entorno, con la Universidad, para mejorar la calidad de la enseñanza. La evaluación al servicio de la educación es una herramienta fundamental si se inserta como tal en los proyectos educativos de los centros.

En esta línea hemos comenzado, en el espacio y en el tiempo de realizar un seguimiento de la implantación de la LOGSE.

Por ello hemos comenzado con el Primer curso del 1^{er} ciclo de E. Primaria.

Las pruebas tendrán una intención de seguimiento individual del progreso de los alumnos en los objetivos de aprendizaje para orientarlos. Sólo los que preparemos para el último proyecto del 2^o curso de Primaria tendrán el propósito de emitir certificación tal y como corresponde.

b) Áreas de contenido. Durante el presente curso hemos podido atender, por razones de tiempo, recursos y expertos, las siguientes áreas:

- Conocimiento del medio
- Lenguaje
- Matemáticas

c) Grupos a medir. Fueron todos los alumnos de doce centros del entorno de Granada que en situaciones ordinarias cursaban el 1^{er} año del 1^{er} ciclo de Primaria.

d) El proyecto requirió un grupo amplio de coordinadores de áreas (3), redactores de ítems (6), colaboradores (15) y expertos en áreas de contenido (12) que coordinados por los autores de este informe trabajamos durante el presente curso, 1992-93.

2. Preparación y revisión de objetivos

Desde una metodología racional, se revisaron los objetivos propuestos por diferentes editoriales para este momento del curso académico, para este nivel y para cada área sí como los documentos legales y técnicos del MEC y publicaciones científicas.

— Las editoriales consultadas fueron: Alhambra-Logman, Anaya, Santillana, S. M. Bruño y Luis Vives.

— Mediante esta revisión se llegó a un consenso sobre un listado de objetivos para cada área. (Ver anexo I).

— A continuación y siguiendo la misma metodología, se estableció (hipotetizó) la relación de cada objetivo con los objetivos de área para el Ciclo y para la etapa así como la conexión con los objetivos generales de educación para el ciclo y etapa (ver anexo II).

3. Planificación adicional de la prueba

a) Selección de objetivos para la prueba. Junto al listado de objetivos antedichos, se describieron los contenidos y las tareas con que más podían relacionarse para presentarlo a los profesores de aula al objeto de que seleccionasen aquellos que coincidían con los programados por ellos, una vez definido su proyecto curricular.

Así se seleccionó un listado de objetivos únicos para todos los centros.

b) Redactores de ítems. Actuaron como tal los coordinadores de área: D^a M^a Teresa Pozo Morate (lenguaje); D. Mariano Machuca Aceituno (lenguaje) y D. Nicolás Morcillo Delgado (Matemáticas). Todos ellos alumnos del programa de doctorado de nuestro Departamento, licenciados en Pedagogía y expertos en cada una de las áreas.

4. Primera redacción de ítems

Una vez establecidos los diferentes conocimientos y habilidades que deberían comprender cada objetivo, éstos se transformaron en tareas (ítems) hipotetizando la relación directa entre cada pregunta y el objetivo que pretendía medir (de la Orden, 1982).

5. Validación de contenido

a) Especialistas. Actuaron como expertos en contenido 12 profesores de EGB con experiencia en Ciclo inicial superior a 5 años y más de 10 en la etapa de EGB. Como expertos en medida actuaron los autores de este informe.

b) Aspectos considerados en cada ítem. Determinar la validez de cada ítem implicaba considerar para cada uno estos tres rasgos.

COHERENCIA o grado con que la formulación del ítem se ajustaba al objetivo que dice medir.

REPRESENTATIVIDAD o grado en que la pregunta formulada es considerada como la más acertada de todas las posibles.

CALIDAD TÉCNICA o grado en que la pregunta expresa con claridad, sin pistas gramaticales falsas, la acción que debe realizar el alumno.

Para cada una de estas categorías se dieron instrucciones precisas a los jueces siguiendo las orientaciones de Hambleton, 1984b.

c) Instrumentación. A cada juez se le suministró:

* Una hoja de respuestas que comprendían:

— N^o de objetivos a medir numerados y formulados

— N^o de ítems que pretendían medir tales objetivos.

— Junto a cada ítem, se acompañaba una escala de valoración de 5 puntos. (ver anexo III).

* Un cuadernillo con la redacción de cada ítem y su numeración correspondiente.

Junto a este material, se dieron instrucciones precisas a los jueces de las que extraemos, como ejemplo, las concernientes a la valoración de COHERENCIA:

«Primero lea cuidadosamente la formulación de cada objetivo, así como, los conceptos y habilidades relacionadas con él. Después por favor, indique hasta que punto usted cree que cada pregunta refleja el dominio que pretende medir. Justifique la pregunta solamente sobre la base de ajuste entre su contenido y el contenido definido por el objetivo que la pregunta pretende medir.

Use la escala de valoración que se acompaña como se muestra a continuación»:

Pobre	Regular	Bueno	Muy Bueno	Excelente
1	2	3	4	5

La forma de expresar su juicio será rodeando con un círculo la valoración que consideres en las escalas que siguen a cada ítem.

d) Análisis de resultados:

Se expresan y a modo ejemplificador «Los obtenidos en el —área de lenguaje—; **Índices de congruencia: ítem-objetivo.**

Para este área, dado que cada objetivo pretendía ser medido por un nº de ítems de forma excluyente, o sea, ningún ítem pretendía medir aspectos de dos o más objetivos, se utilizó la formulación propuesta por Hambleton, 1984:

$$I_{ik} = \frac{(N - 1) \sum_{j=1}^R X_{ijk} + N \sum_{j=1}^R X_{ijk} - \sum_{j=1}^R X_{ijk}}{2 (N - 1) n}$$

donde

I_{ik} = índice de congruencia ítem-objetivos

N = nº de jueces

X_{ijk} = Evaluación del juez j para el ítem K en el objetivo i .

Para ello hubo de reducirse a tres todas las valoraciones. -1, 0 y +1, coincidiendo 0 con la posición central (3); -1 para las valoraciones 1 y 2 y +1 para las valoraciones 4 y 5.

Se recogieron en tablas las evaluaciones de los jueces y los índices de congruencia ítem-objetivo hallados para cada uno de los aspectos evaluados.

A continuación fijamos los criterios de adecuación en función de los I_{ik} obtenidos.

Como *criterio absoluto* se fijó a priori el valor de 0'75 (que se considera un índice bastante adecuado para 4 jueces). El criterio relativo se calculó, para cada uno de los aspectos evaluados hallando la media de los índices de congruencia obtenidos en cada uno de ellos por lo que es diferente para cada aspecto enjuiciado, resultando los siguientes valores:

- Coherencia 0'92
- Representatividad 0'8
- Calidad técnica 0'82

b) Experiencia de resultados. En tablas (una para cada área y aspecto evaluado) se recogen las evaluaciones de los jueces y los índices de congruencia ítem objetivo hallados para cada uno de los tres aspectos, colocando un asterisco junto los índices que sobrepasan los valores de los criterios absoluto y relativo (ver anexo IV).

Para tener una visión global de la evaluación de cada ítem recogimos en otra tabla sólo los índices de congruencia de cada uno de los tres aspectos evaluados, (ver anexo V).

Es conveniente que los ítems clasificados según los resultados obtenidos se especifiquen en un cuadro en el que aparezcan los ítems ACEPTADOS, que se admiten sin modificación salvo aquellos que van acompañados de algunos códigos, que en nuestro caso (ver anexo VI) hemos establecido con:

CT: Indica que se debe revisar la calidad técnica del mismo pues ha obtenido en este aspecto un índice excesivamente bajo.

R. CT.: Indicaría que se debe revisar el porqué no se le considera realmente representativo del objetivo con el que se asocia y también la calidad técnica:

En aquellos ítems, en que el Índice de Representatividad es bastante bajo, los hemos aceptado siempre que en Coherencia y Calidad Técnica tengan índices altos. En estos casos partimos del supuesto de que no es relevante que no se vea a este ítem como uno de los más representativos para cubrir un objetivo.

En la misma tabla (ver anexo VI) deben aparecer los ítems POSIBLES y los RECHAZADOS.

Englobamos dentro del grupo de posibles aquellos ítems que, en general, en los tres índices de congruencia no tienen una puntuación inferior a 0'25 del criterio más bajo para cada aspecto, (aquí el

criterio absoluto). En estos casos es necesario revisar estos ítems para observar donde se encuentra el motivo de una baja valoración, teniendo alguno de ellos una puntuación alta en algún índice.

Para las áreas de **matemáticas y conocimiento del medio** se aportan sólo las variaciones metodológicas en el análisis de los enjuiciamientos.

Tales variaciones son producidas por los expertos en contenido y medida al hipotetizar para cada ítem, relación de medida con varios objetivos.

En efecto los especialistas han considerado que tal y como han redactado los ítems, éstos pueden medir aspectos de varios objetivos.

Para este caso utilizamos para el cálculo del índice de congruencia ítem-objetivo la formulación propuesta por Rovinelli y Hambleton, 1977

$$I_{jk} = \frac{N}{2N-2} (\mu_k - \mu)$$

donde:

I_{jk} = índice de congruencia ítem-objetivos.

N = nº de objetivos totales evaluados por los jueces.

μ_k = media de valoración del ítem en el objetivo k.

μ = media del ítem en todos los objetivos.

Esta formulación está pensada para aquellas situaciones en que se presenta a los jueces una lista de objetivos y otra de ítems sin hipotetizar relaciones entre ellos «a priori». Como en nuestro caso se han presentado ya relacionadas, hemos tenido que partir del supuesto de que todos los jueces evaluarían como -1 aquellos ítem de cuya formulación no se desprende relación alguna con los objetivos que pretenden medir.

6. REVISIÓN DE ÍTEMS

Los datos procedentes de la anterior validación fueron remitidos a los coordinadores de área que a su vez fueron los escritores de ítems para que, tras tener en cuenta el análisis, procedieron a una segunda redacción.

7. EXAMEN DE CAMPO DE LOS ÍTEMS Y ENSAMBLAJE DE LA PRUEBA

Dado el propósito de la prueba, se hizo coincidir con la administración de la misma a los alumnos con el examen de campo.

Ambas situaciones se están desarrollando en el momento de escribir esta comunicación por lo que no nos queda sino especificar el resto de acciones que quedan por realizar para dar por terminada la prueba.

Respecto al examen de campo, será utilizado para detectar preguntas defectuosas, comprobar o chequear sesgos, determinar índices de dificultad y consistencia de puntuaciones o decisiones.

Respecto al ensamblaje de la prueba, determinaremos la longitud total y el nº de preguntas de la prueba por objetivo. Prepararemos la composición, los cuadernos de respuestas de los alumnos y las claves de puntuación.

Esperamos, si es aceptada nuestra comunicación, aportar en su momento toda la información complementaria al proceso llevado hasta el momento.

REFERENCIAS BIBLIOGRÁFICAS

- ARTACHO LÓPEZ, R. (1992): *Elaboración del Proyecto Curricular E.P.* Zaragoza. Luis Vives.
- DEPARTAMENTO DE PROYECTOS DEL GRUPO ANAYA (1992): *Proyecto Curricular para la Educación Primaria-Andalucía*. Madrid. Anaya.
- EQUIPO ANAYA (1992): *Proyecto Las Estaciones*. Madrid. Santillana.
- GLASER, R. (1963): Instructional Technology and the measurement of Learning outcomes. *American Psychologist*, 18, 519-521.
- GRAY, W. M. (1978): A comparison of Piagetian theory and criterion-referenced measurement. *Review of Educational Research*, 48, 223-249.
- GRUPO S. M. (1992): *Proyecto Curricular y Programación*. Madrid. S.M.
- HAMBLETON, R. (1990): Criterion referenced testing. Methods and practices. En GUTKIN, T. y REYNOLDS, C. (Eds.): *The handbook of School Psychology*. New York. Wiley.
- HAMBLETON, R. K. (1984): Validating the test scores. En R. A. BERK (Ed.): *A Guide to Criterion-referenced Test Construction*, 2ª Ed. Baltimore, The Johns Hopkins University Press.
- HAMBLETON, R. K. y SWAMINATHAN, H. (1985): *Item Response Theory. Principles and Applications*. Boston. Kluwer-Nijhoff.
- JIMÉNEZ MORENO, E. (1992): *Proyecto Albanta*. Madrid. Alhambra Logman.
- JUNTA DE ANDALUCÍA. BOJA. 20-6-92: *Decreto 105/1992 del 9 de Junio*. Sevilla.
- MEC. BOE 26-6-91: *Decreto 1007/1991 del 14 de junio*. Madrid.
- ORDEN HOZ, A. (1982): *La evaluación educativa*. Docencia. Buenos Aires.
- PÉREZ JUSTE, R. (1992): *El modelo evaluativo de la Reforma Educativa*. Madrid. UNED.
- POPHAM, W. J. y HUSEK, T. R. (1969): Implications of criterion-referenced measurement. *Journal of Educational Measurement*, 6, 1-9.
- RIVAS, F. y ALLANTUD (1989): *Evaluación criterial en Enseñanza Primaria*. Madrid. CIDE.
- ROVINELLI, R. J. y HAMBLETON, R. K. (1977): On the use content specialist in the assessment of criterion-referenced test item validity. *Dutch Journal of Educational Research*, 2, 49-60.

PROCESO DE ELABORACIÓN DE UN INSTRUMENTO DIAGNÓSTICO DE LA EXPERIENCIA DOCENTE. (ESTUDIO EXPLORATORIO)

Concepción Riera Quintana

Universidad de La Laguna. Tenerife. Islas Canarias

La experiencia docente es un constructo que ha sido utilizado en muchas investigaciones de distinta índole como variable predictora, diferenciadora del profesorado: estudios sobre la naturaleza y adquisición del conocimiento, estudios centrados en la formación del profesorado, que intentan describir y explicar las diferencias entre profesores noveles y expertos con el fin de extraer conclusiones cara a los programas educativos y las estrategias de enseñanza...

Sin embargo, la definición del término «experiencia docente», así como su utilización como variable clasificatoria de grupos de profesores no presenta la claridad que sería deseable para el avance del conjunto de la investigación educativa. A veces se presenta la experiencia —único término utilizado hasta la fecha en castellano— como el conocimiento acumulado durante el ejercicio de la docencia (experience); otras como la habilidad de los profesionales para realizar su trabajo (expertise); e incluso, como eficacia del profesorado, medida generalmente en términos de resultados en el alumno, (proeficiency). Pero no vamos a entrar en el problema de indefinición del concepto, ya que hay autores que se han dedicado a ello con detenimiento (Lampert y Clarck, 1990). Sólo señalar que para poder medir un concepto primero hay que definir claramente y de forma unívoca lo que se está midiendo.

En cuanto a la utilización de la experiencia docente como variable, existe desde 1986 hasta la fecha una gran cantidad de investigaciones publicadas que basan sus resultados en el estudio diferencial de profesores expertos y novatos, e incluso hay revisiones hechas sobre los mismos (Reynolds, 1992; Kagan, 1992). En todas se pueden apreciar problemas comunes en el uso de la experiencia docente como variable:

- Se valora casi exclusivamente en función del número de años de ejercicio.
- Se confunden los límites temporales que separan a un profesor experto de un novato.
- Hay confusión también en cuanto a los términos —que se usan de forma intercambiable— para denominar a diversos tipos de profesionales (novatos, principiantes avanzados, postulantes...).
- A nivel práctico, sólo se consideran los valores extremos de la variable dicotómica, sin posibilidad de gradaciones intermedias.

Todo lo anterior genera, entre otras, la dificultad de generalización de los resultados obtenidos en las investigaciones particulares y, por tanto la imposibilidad de hacer uso práctico de las conclusiones extraídas.

La presente comunicación pretende, a grandes rasgos, describir el proceso seguido en la fase exploratoria de construcción de un instrumento diagnóstico que permita una mejor clasificación del profesorado a la hora de realizar cualquier tipo de estudios en el que la experiencia sea una variable de interés. Dos objetivos han guiado el trabajo:

— Profundizar en la estructura interna del constructo «experiencia docente» a fin de poder aclarar lo que se quiere medir.

— Dar pautas para la elaboración del instrumento que permita valorar a los profesores con diversos grados de experiencia.

A continuación presentamos los pasos sucesivos y los resultados obtenidos hasta el momento.

1. EL TORBELLINO DE IDEAS

La utilización de esta técnica surgió, una vez revisada la literatura disponible, como necesidad de obtener información directa de los propios profesionales en distintos ámbitos educativos. En las sesiones realizadas se les pedía a los participantes que expresaran sus ideas sobre la experiencia docente y los rasgos que definen a los expertos.

Las sesiones fueron grabadas en cassette. Después de transcribir las cintas se procedió a lo siguiente:

— Realizar una primera selección de las ideas, eliminando aquellas demasiado ambiguas, reiterativas o poco relacionadas con el concepto a estudiar. Se obtuvieron de esta forma un total de 132 frases.

— En función de las frases seleccionadas, definir las dimensiones que delimitan el constructo experiencia docente.

Inicialmente fueron detectadas un total de nueve dimensiones:

años de ejercicio, nivel de reflexión teórica sobre la práctica, habilidades personales, dominio del contenido, formación inicial, «estrategias de navegación» en el centro, formación permanente, habilidades pedagógicas y actitud hacia la innovación.

— Adjudicar cada frase a una dimensión. Hay que señalar que no todas las dimensiones resultaron con el mismo número de frases. En concreto la dimensión «reflexión teórica sobre la práctica» fue eliminada por sólo clasificarse en ella dos de las frases recogidas.

2. SISTEMA DE JUECES

Esta técnica se utilizó con dos finalidades distintas:

a) Revisar la clasificación de las frases en las diferentes dimensiones hecha por los investigadores.

b) Seleccionar el conjunto de frases más representativo de cada dimensión.

Para ello se solicitó a seis jueces —psicólogos y pedagogos— que realizaran de forma individual la siguiente tarea: una vez presentada cada dimensión en una tarjeta y explicado el significado de la misma, se entregaban de forma desordenada las frases en tarjetas. Ellos debían situarlas en la dimensión que considerasen oportuna. Se les especificó que cada frase podía ser representativa de cualquiera de las dimensiones o de ninguna, e incluso podía haber frases situadas en más de una dimensión. También se les pidió a los jueces que aportaran todas las sugerencias que considerasen interesantes para la investigación.

Los datos de cada sujeto fueron recogidos en una hoja individual y vaciados en una tabla de doble entrada; en las filas constaban las dimensiones y en las columnas, de acuerdo con la clasificación de los investigadores, la ubicación correcta o incorrecta (intra y extra dimensión) de las frases dependiendo del número de acuerdos interjueces.

Como resultado del análisis de datos realizado se eliminaron aquellas frases que eran situadas en más de una dimensión o fuera de las dimensiones contempladas. Se seleccionaron aquellas frase en que había total acuerdo interjueces y se resituaron frases consideradas por los investigadores como propias de una dimensión y clasificadas por todos los jueces en otra diferente. Fueron seleccionadas 51 frases.

Atendiendo a las sugerencias realizadas por los jueces y a los datos obtenidos se realizó una nueva estructuración de las dimensiones, que quedaron reducidas a seis: n.º de años de ejercicio, habilidades personales, dominio del contenido, Formación pedagógica, estrategias de navegación/colaboración entre profesores y habilidades pedagógicas.

3. ESCALAMIENTO MULTIDIMENSIONAL. (MDS)

Con el fin de esclarecer la estructura interna del concepto «experiencia docente» y determinar los aspectos a tener en cuenta en su medición aplicamos un cuestionario cuyos datos fueron analizados mediante un escalamiento multidimensional através del programa informático INDSCAL.

Para la construcción del cuestionario se eligieron 24 de las 51 frases seleccionadas en la fase anterior (4 frases por cada dimensión). Esta reducción se debió a limitaciones del programa¹: el cuestionario fue aplicado a un total de 31 profesores. El trabajo que debían realizar consistía en comparar las 24 frases entre sí, valorandola semejanza de las ideas contenidas en las frases con una escala de similitud (desde 1, nada parecidas, hasta 7, iguales).

Una vez aplicado el programa se escogió la solución de dos dimensiones. Cada dimensión o vector, permite identificar los criterios en función de los cuales los sujetos valoran las semejanzas entre las frases y, por tanto, los criterios por los que valoran la experiencia docente.

El primer vector ha sido interpretado como «Habilidades y capacidades profesionales», con una dimensión personal y otra colectiva. El segundo vector se definió como «Conocimiento pedagógico», referido tanto al conocimiento didáctico como del alumnado y la materia.

Analizando el agrupamiento de los frases en torno a estos dos vectores, obtenemos tres grupos diferenciados:

a) Un primer grupo se aglutina en el cuadrante inferior izquierdo del gráfico y hace referencia a la experiencia docente como un conjunto de habilidades y capacidades profesionales desarrolladas colectivamente en el ejercicio de la docencia.

b) El segundo grupo forma un racimo distribuido por los dos cuadrantes superiores y recalca los aspectos referidos al conocimiento del alumno y de las estrategias didácticas necesarias para desarrollar en él los aprendizajes.

c) El tercer grupo de frases se reúne en torno al primer vector, en su extremo derecho y acentúa los aspectos individuales o personales de la experiencia docente en lo que se refiere a la adquisición del conocimiento pedagógico.

A modo de conclusión de este estudio:

— Existen diferentes concepciones sobre lo que es experiencia docente. Si bien estas concepciones no son contradictorias entre sí, ninguna contempla el n.º de años ejerciendo la docencia como el único factor determinante del nivel de experiencia de los profesores.

— La experiencia docente es un constructo multidimensional. Su medición requerirá el uso de distintos instrumentos y metodologías que permitan acceder a cada uno de los aspectos aquí referidos (escalas de actitud, observaciones...).

1 $N \times E^2 < 18000$. N: sujetos. E: n.º de estímulos (frases).

REFERENCIAS BIBLIOGRÁFICAS

- FLODEN, R. E. y KLINZING, H. G. (1990): What Can Research on Teacher Thinking Contribute To Teacher Preparation? A Second Opinion. *Educational Researcher*, 4, 15-20.
- HUBERMAN, M. (1990): Las fases de la profesión docente. Ensayo de descripción y previsión. *Curriculum*, 2, 139-160.
- KAGAN, D. M. (1992): Professional Growth Among Preservice and Beginning Teachers. *Review of Educational Researcher*. 2, 129-169.
- LAMPERT, M. y CLARK, C. M. (1990): Expert Knowledge and Expert Thinking in Teaching: A Response To Foden And Kinzing. *Educational Researcher*. 5, 21-23.
- MARCELO, C. (1987): *El pensamieto del profesor*. Barcelona: CEAC.
- PEARCE, Ph. L. y AMATO, P. R. (1980): A Taxonomy of Helping: a Multidimensional Scaling Analysis. *Social Psychology Quarterly*. 4, 363-371.
- POZO, J. I. (1989): *Teorías cognitivas del aprendizaje*. Madrid: Morata.
- PRADO, D. (1986): *El torbellino de ideas. Hacia una enseñanza más participativa*. Madrid: Cincel.

DISEÑO DE UN INSTRUMENTO DE ANÁLISIS DE LA ENSEÑANZA COMUNICATIVA DE LENGUAS EXTRANJERAS

Mercedes Verdú Jordá,
Universidad de Murcia
Enrique Iglesias Verdegay
Universidad de Murcia

INTRODUCCIÓN

El enfoque comunicativo de la enseñanza de lenguas extranjeras ha dado un giro total a los roles que tanto el profesor como el alumno tienen que desempeñar en el aula. Con el término *roles* queremos indicar tanto el papel que asumen los alumnos y profesor durante el desarrollo de las tareas y actividades como el de las relaciones sociales e interpersonales que se establecen entre los participantes (Breen y Candlin, 1980; Nunan, 1985).

La evaluación de los programas de enseñanza de lenguas extranjeras es una necesidad imperiosa que tienen los responsables de los centros educativos y ésta ya no puede basarse en la medición de la diferencia entre el estado inicial y final de los alumnos pues, como hemos apuntado anteriormente, existen toda una serie de procesos que generan cualidades que los instrumentos puramente cuantitativos no nos pueden reflejar.

En los años 60 la evaluación de los programas de enseñanza de lenguas extranjeras tenía como finalidad determinar cuál era el **método** correcto, basándose en la suposición de que lo que se aprendía estaba totalmente determinado por el tipo de método que se aplicaba. Los estudios de Sherer y Wertheimer (1964) y los de Smith (1970) pusieron de manifiesto la esterilidad de las investigaciones basadas en percepciones metodológicas globales y concluyeron que no tiene sentido probar que un método es superior a sus competidores. Según Politzer (1970: 43) *la gran complejidad del proceso de enseñanza hace muy difícil hablar en términos absolutos sobre mecanismos de enseñanza «buenos» y «malos»*.

Estos resultados, a nuestro juicio, deberían llevarnos a replantear el tipo de investigación que podemos desarrollar en el aula. Consideramos que se debe pasar de una aproximación prescriptiva a una *descriptiva* y además pasar de las técnicas a los *procesos*. Ya no es posible asumir que todo lo que sucede es debido a un método o conjunto de técnicas que se llevan a la práctica. En estos momentos se empieza a sentir que algo por debajo del nivel de la técnica (más interactivo y menos pedagógico) está teniendo lugar a cada momento y que el estudio de las interacciones que se dan en el aula puede ser un campo fructífero (Allwright, 1988).

Para nosotros el concepto de interacción no se refiere únicamente a los intercambios que implican una comunicación auténtica sino que abarca además todos los intercambios orales que ocurren en el

aula, incluidos los que se dan durante el desarrollo de actividades centradas en la forma de la lengua (Allwright, 1984). Según Jarvis (1968), un investigador pionero en el análisis de las interacciones, es preciso diseñar un sistema de observación que refleje las características especiales del aula de lengua teniendo en cuenta que la lengua se usa en el aula tanto con propósitos reales como con propósitos de aprendizaje por medio de ejercicios controlados (drills).

ANÁLISIS TRADICIONAL DE LA INTERACCIÓN EN EL AULA

A partir de los años 60 se crean numerosos instrumentos de observación de las interacciones en el aula que se centran, en primer lugar, en analizar el tipo de dependencia que se establece entre los comportamientos del alumno con respecto al ambiente del aula y la interacción que genera el profesor y en segundo lugar, no se persigue un análisis cuantitativo, aunque utilicen medidas de frecuencias y duraciones de tiempo. En esta línea deberemos citar el sistema FLINT (Foreign Language Interaction) de Moskowitz (1976) que podemos clasificar en el sistema de signos (Rosenshine y Furst, 1973). Las categorías que se contemplan en este instrumento se usan con dos fines específicos: como herramienta de investigación, para intentar averiguar qué constituye una *buena* enseñanza de la lengua y como herramienta para suministrar información de reflexión en la formación de profesores. Long (1980) critica este esquema de observación poniendo en duda la idoneidad de basarse en categorías que fueron pensadas por Flanders (1960) para otras áreas del curriculum. Por otra parte, Leona Bailey (1975) considera que las categorías de comportamiento de Moskowitz están abiertas a múltiples interpretaciones por parte de distintos observadores, dada la ambigüedad de las mismas, lo que entraña bastante dificultad para su codificación en tan corto período de tiempo (tres segundos).

Fanselow (1977) construyó el esquema FOCUS (Foci for Observing Communications Used in Settings) que utiliza como unidad de análisis el *turno* (move) del discurso pedagógico y las categorías de la dimensión del propósito pedagógico. La objeción más seria que se le hace a este esquema de observación es que es un sistema altamente complejo, infinitamente mucho más complejo que el FLINT y por tanto de difícil aplicación para la observación práctica (Allwright, 1988). En esta misma corriente de análisis de la interacción se deben destacar el esquema de Naiman, Frölich, Stern y Todesco (1978) que aporta una descomposición más detallada de las unidades lingüísticas a analizar pero que sigue manteniendo como unidad de análisis el *turno* pedagógico.

Estos esquemas de observación, además de no conservar la estructura general de la intervención observada y ser de una fiabilidad dudosa, se centran en lo que dice el profesor, ignorando la comunicación no verbal (Long, 1980). Como consecuencia de estas deficiencias a partir de los años 80 se va a producir múltiples aportaciones provenientes de diversas áreas relacionadas con el aprendizaje, así como por los resultados de los estudios empíricos a pequeña escala dentro del aula. Se deben citar las aportaciones de Mitchell, Parkinson y Johnstone (1981) que proponen una serie de criterios por los que se debe guiar quien quiera diseñar un instrumento de análisis sistemático de las lecciones. Estos criterios son los siguientes:

a) Cualquier sistema debe apoyarse en la *interpretación teórica en uso* sobre el proceso de aprendizaje de las lenguas extranjeras.

b) Los sistemas deberán permitir la *codificación multidimensional del discurso*. Es decir, deben incluirse tantas dimensiones como parezcan significativas sobre la base de la interpretación teórica anteriormente mencionada.

c) Al aceptar el concepto de discurso como un fenómeno con su propio estructura interna, los sistemas de análisis deben intentar mantener esta estructura tanto como sea posible, *adoptando unidades de discurso*, a uno o más niveles, *como unidades básicas de análisis*, más que unidades basadas en tiempo o unidades no analíticas.

A partir de estos criterios Mitchel et al. diseñaron su instrumento de observación que utilizaron en un estudio proceso-producto. La principal novedad que introdujeron fue la decisión de adoptar lo que denominamos *segmento* como unidad básica de análisis. Como ellos mismos exponen:

«En este caso, se dio prioridad al análisis de la enseñanza de la LE en función de unidades «naturales» relativamente amplias, que se corresponden aproximadamente con el concepto práctico de los profesores de «actividades de enseñanza» (tales como «ejercicios estructurales», «repetición», etc.).» (Mitchell et al., 1981:11)

En su análisis final Mitchell et al. (1981) demuestran que existe un vacío entre las condiciones óptimas para el aprendizaje de LE en las aulas y la práctica de aula observada en este estudio.

Posteriormente Ullman y Geva (1984) diseñaron su esquema denominado TALOS (Target Language Observation Scheme) de una perspectiva mucho más amplia que sus antecesores. Este esquema contiene dos secciones bien diferenciadas: una sección de *categorías de baja inferencia* para su codificación en tiempo real y otra sección de *categorías de alta inferencia*.

Finalmente hemos de mencionar el esquema de observación denominado COLT (Communicative Orientation of Language Teaching) que fue originalmente diseñado por Allen, Fröhlich y Spada (1984) que consta de dos partes. La primera consta de cinco parámetros principales diseñados para medir hasta qué punto un determinado tipo de implementación puede considerarse comunicativa y considera como unidad básica de análisis la *actividad*. La segunda analiza las actividades a nivel de interacción verbal.

Los instrumentos de observación anteriormente reseñados coinciden en la necesidad de identificar los participantes individuales de la interacción y en incluir el tópico o contenido, aunque la descripción individual de esta categoría difiera considerablemente. Tres esquemas incluyen la categoría *actividad* y el uso de *destrezas* y *materiales*. Según Chaudron (1988) esta falta de coincidencia en la selección de las categorías ilustra que ninguno de los esquemas incluye todas las dimensiones de información sobre la interacción del aula potencialmente relevantes, lo que no invalida la utilidad de los esquemas, pero sigue argumentando que cuando los investigadores que intentan observar las mismas dimensiones básicas no seleccionan las mismas categorías de análisis, sucede que no sólo los resultados no son comparables sino que, al menos uno de los esquemas, si no todos, no están empleando un conjunto válido de categorías de observación.

NUESTRO ESQUEMA DE OBSERVACIÓN

Para su diseño partimos del esquema COLT (Allen, Fröhlich y Spada, 1984) del que seleccionamos las dimensiones más pertinentes para nuestro trabajo y, posteriormente, completamos el esquema con otras nuevas que pudieran medir todos los rasgos específicos que posiblemente se iban a dar en las aulas objeto de estudio.

Las dimensiones incluidas derivan de las teorías sobre la *competencia comunicativa* (Canale y Swain, 1980; Breen y Candlin, 1980) y de la *enseñanza comunicativa* de la lengua (Brumfit, 1980; Johnson y Morrow, 1981; Widdowson, 1978, 1979; Breen, 1987), así como de la *investigación sobre la adquisición* de una segunda lengua extranjera en el aula (Ellis, 1990, Chaudron, 1988, van Lier, 1988) y de la *observación llevada en el aula* (Allwright, 1984, 1988; Fanselow, 1977; Mitchell, Parkinson y Johnstone, 1981; Ullman y Geva, 1984; Allen et al., 1984) que sugieren una serie de factores que pueden influir en el proceso de aprendizaje de una lengua extranjera.

El esquema que hemos confeccionado consta de doce categorías o dimensiones generales en las que se recoge información sobre diversos factores categorizados en una serie de subdimensiones. Estos

factores dentro de cada categoría aparecen en nuestro esquema en un orden de mayor a menor grado de comunicatividad siendo los situados en último lugar los que representan los rasgos más innovadores. El esquema de observación es el que queda reflejado en la página siguiente.

La unidad básica de análisis es la *actividad* y en torno a ella se han analizado las dimensiones de la observación.

SISTEMA DE ACTIVIDADES

Se ha seleccionado la actividad como unidad de análisis porque de acuerdo con Shavelson y Stern (1981:477) «las actividades y no el modelo prescriptivo son el centro de la planificación del profesor...». De acuerdo con Crookes (1990) la unidad de análisis ha de ser fácilmente identificable y ha de tener una realidad psicológica como unidad de comportamiento. Una posible definición de actividad sería la siguiente:

La mayor unidad de segmentación de un evento en una clase de lenguas distinguible sobre la base de los procesos lingüísticos-cognitivos y el propósito pedagógico que ambos implican.

En la revisión bibliográfica que hemos realizado hemos detectado que no existe un sistema de clasificación de las actividades utilizadas en cada uno de los estudios, lo que constituye una pérdida de información lamentable ya que no puede haber control sobre el mayor o menor grado de comunicatividad dentro del conjunto de actividades que se dan en la clase de lengua extranjera.

En la realización de nuestro sistema de actividades hemos prestado especial atención a dos aspectos: uno ha sido el de ofrecer un **conjunto completo de descriptores** de las actividades de forma que sean fácilmente identificables y el otro ha sido el de asegurarnos que se puedan **codificar de forma fiable**. Aplicando la kappa de Cohen los resultados que hemos obtenido en la codificación del sistema de actividades hemos registrado un acuerdo entre codificadores que oscilaba entre el 82% y el 88%, según el tipo de actividades. La lista completa de actividades puede consultarse en Verdú (1993) donde se expone detalladamente la construcción de las mismas.

PROCEDIMIENTO

A continuación pasamos a describir sucintamente el procedimiento que hemos seguido para llevar a cabo nuestro estudio.

— Sujetos del estudio.

El esquema de observación y el sistema de actividades se han aplicado en nueve aulas de 6º, 7º y 8º de EGB cuyo profesores habían participado en un curso de perfeccionamiento de Inglés organizado por la Subdirección General de Perfeccionamiento del Profesorado, durante el curso 1985/86. Los centros que han participado en la experiencia diferían en cuanto a su ubicación, status socioeconómico y actitudes ante la enseñanza. El número de alumnos por aula oscilaba entre 20 y 30.

— Obtención de datos.

Los datos se obtuvieron durante el curso 1987/88 y finalizó en el curso 1988/89 observando nueve profesores durante tres sesiones consecutivas —equivalentes al desarrollo de una unidad— de aproximadamente 60 minutos cada una y siempre sobre los mismos alumnos.

Todas las sesiones fueron grabadas en audio y vídeo para ser contrastadas con las observaciones efectuadas en la clase utilizando el instrumento de observación y con codificación en tiempo real mediante, al menos, dos observadores.

Las grabaciones fueron codificadas de nuevo por otros observadores y no se encontraron diferen-

cias significativas entre las codificaciones realizadas en tiempo real y las codificadas posteriormente. Esto prueba la estabilidad y fiabilidad de nuestro instrumento de observación.

— **Análisis de los datos.**

Los datos recogidos mediante el esquema de observación fueron sometidos a los siguientes análisis:

a) En primer lugar se hicieron los cálculos de porcentajes de tiempo dedicado a cada actividad en cada una de las dimensiones del esquema de observación.

b) Cálculo del índice de fiabilidad de las observaciones efectuadas por los observadores obteniéndose una alfa de Cronbach igual a 0,99 en las dimensiones observadas.

c) Se ha hecho un análisis factorial de cada una de las dimensiones del total de 698 actividades observadas obteniéndose **cinco factores** que explican el 74,8% de la varianza. Los resultados de este análisis pueden resumirse de esta forma:

Factor 1: Dimensiones	Carga	Denominación de los factores
Uso Lengua del Alumno	.94	
Comprensibilidad	.90	
Información alumno predecible	.87	Interacción verbal del alumno
Longitud discurso del alumno	.87	
Información alumno genuina	.85	
Feedback	.70	
Factor 2: Dimensiones	Carga:	
Información Profesor Predecible	.93	
Información Profesor Genuina	.90	
Longitud Discurso del Profesor	.82	Interacción verbal del Profesor
Uso de Lengua del Profesor	.82	
Autenticidad de Uso de Lengua	.59	
Tópico	.49	
Longitud del Texto	.43	
Factor 3: Dimensiones	Carga	
Grado de Control de la Actividad	.82	
Control sobre Turno del Alumno	.77	
Control sobre el Tópico	.67	Control de la actividad
Grado de interacción	.64	
Factor 4: Dimensiones	Carga:	
Relación con Lengua 1	.85	Relación lengua 1
Factor 5: Dimensiones	Carga	
Contextualización	.87	Contextualización

Con este análisis factorial se ve la coherencia de nuestro esquema de observación con el enfoque teórico que sustenta nuestra visión de la enseñanza de las lenguas: el enfoque comunicativo.

d) Estudio de las dimensiones que se dan en cada una de las fases de una clase comunicativa: motivación, input, práctica y aplicación o transferencia. Nuestro esquema de observación nos permite detectar que existen diferencias entre las dimensiones de las actividades desarrolladas entre las diferentes fases.

RESULTADOS

Los instrumentos diseñados: esquema de observación y sistema de actividades nos permiten detectar el nivel de comunicatividad que se está dando en las clases así como determinar la **secuencia** que han seguido los profesores en la elaboración de sus unidades didácticas.

El esquema de observación nos permite detectar en qué aspectos los profesores han sido innovadores, pudiéndose establecer un perfil innovador para cada uno de los profesores y determinar qué aspectos deberían ser modificados para ampliar la innovación en extensión y profundidad.

Esperamos que estas herramientas puedan llegar a ser instrumentos útiles para que los profesores pudieran reflexionar sobre una serie de cuestiones relacionadas con las ventajas y desventajas de su aplicación de un enfoque comunicativo frente a las que ofrece la aplicación de una metodología más estructural y controlada en la enseñanza de las lenguas extranjeras.

BIBLIOGRAFÍA

- ALLEN, J. P., FRÖHLICH, M. y SPADA, N. (1984): The communicative orientation of language teaching: an observation scheme. En J. HANDSCOMBE, R. A. OREM y B. P. TAYLOR (Eds.): *On TESOL '83: The Question of Control* (pp. 231-252). Washington, D.C.: TESOL.
- ALLWRIGHT, R. (1984): The importance of interaction in classroom language learning. *Applied Linguistics*, 5, 156-171
- ALLWRIGHT, R. (1988): *Observation in the Language Classroom*. Londres: Longman.
- BAILEY, L. G. (1975): An observational method in the foreign language classroom: a closer look at interaction analysis. *Foreign Language Annals*, 8(4), 335-344.
- BREEN, M. P. Y CANDLIN, C. (1980): The essentials of communicative curriculum in language teaching. *Applied linguistics*, 1(2), 98-112.
- BRUMFIT, C. J. (1980): From defining to designing: communicative specifications versus communicative methodology in foreign language teaching. *Studies in Second Language Acquisition*, 3(1), 1-9.
- CANALE, C. y SWAIN, M. (1980): Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1, 1-47.
- CHAUDRON, C. (1988): *Second Language Classrooms: Research on Teaching and Learning*. Cambridge: Cambridge University Press.
- ELLIS, R. (1990): *Instructed Second Language Acquisition: learning in the classroom*. Cambridge, MA: Basil Blackwell.
- FANSELOW, J. F. (1977): Beyond Rashomon-Conceptualizing and describing the teaching act. *TESOL Quarterly*, 11(1), 17-39.
- FLANDERS, N. A. (1960): *Interaction Analysis in the Classroom: A Manual for Observers*. Ann Arbor, MI: University of Michigan Press.
- JARVIS, G. A. (1968): A behavioral observation system for classroom foreign language skill acquisition activities. *Mosern Language Journal*, 52, 335-341.
- JOHNSON, K. y MORROW, K. (Eds.) (1981): *Communication in the Classroom*. London: Longman.
- LONG, M. H. (1980): Inside the «black box»: methodological issues in research on language teaching and learning. *Language Learning*, 30, 1-42.
- MITCHELL, R., PARKINSON, B. y JOHNSTONE, R. (1981): *The Foreign Language Classroom: An observational study*, Stirling, Scotland: Stirling Educational Monographs, nº 9, Universidad de Stirling.
- MOSKOWITZ, G. (1976): The classroom interaction of outstanding foreign language teachers. *Foreign Language Annals*, 9, 135-143; 146-157.

- NAIMAN, N., FRÖLICH, M., STERN, H. H. y TODESCO, A. (1978): *The good language learner*. Toronto: O.I.S.E.
- NUNAN, D. (1991): Communicative Tasks and the Language Curriculum. *TESOL Quarterly*, 25, 279-295.
- ROSENSHINE, B. y FURST, N. (1973): The use of direct observation to study teaching. En R. TRAVERS (Ed.): *Second Handbook of Research of Teaching* (pp. 122-183). Chicago: Rand McNally.
- SHAVELSON, R. y STERN, P. (1981): Research on teachers» pedagogical thoughts, judgements, decision and behaviour. *Review of Educational Research*, 51(4), 455-498.
- SCHERER, A. y WERTHEIMER, M. (1964): *A Psycholinguistic Experiment in Foreign Language Teaching*. Nueva York: McGraw-Hill.
- SMITH, P. D. (1970): *A comparison of the cognitive and audiolingual approaches to foreign language instruction: the Pennsylvania foreign language project*. Philadelphia: Center for Curriculum Development.
- ULLMANN, R. y GEVA, E. (1984): Approches to observation in second language classes. En J.B.P. ALLEN y M. SWAIN (Eds.): *Language Issues and Education Policies*. Oxford: Pergamon Press.
- VERDÚ, M. (1993): *Diseño de un instrumento de análisis de la enseñanza comunicativa de lenguas extranjeras*. Universidad de Murcia. Tesis doctoral no publicada.
- VAN LIER, L. (1988). *The classroom and the language learner*. Londres: Longman.
- WIDDOWSON, H. G. (1978): *Teaching Language as Communication*. Oxford: Oxford University Press.
- WIDDOWSON, H. G. (1979): *Explorations in Applied Linguistics*. Londres: Oxford University Press.

EVALUACIÓN DINÁMICA DE LOS PROCESOS Y ESTRATEGIAS COGNITIVAS EN NIÑOS DEFICIENTES MENTALES

Santiago Molina García y Ana Arráiz Pérez
Universidad de Zaragoza

1. JUSTIFICACIÓN

Este modelo ha surgido como alternativa al modelo tradicional, basado en la aplicación de tests psicométricos, de forma totalmente rígida, de los que sólo es posible extraer un C.I. aplicando una metodología cuantitativa.

Missiuna y Samuels (1988, 2) citan las tres áreas en las que hay una coincidencia clara para cuestionar la evaluación normativa, y dos de las cuales están precisamente referidas al proceso de estandarización: inadecuación técnica de los instrumentos e inexperiencia profesional en su administración, inadecuación de los instrumentos estandarizados para la evaluación de los niños cuya aculturación difiere del grupo estandarizado.

Haywood (1977, 12) afirma que la mayor utilidad de este enfoque sería proporcionar información para tomar decisiones políticas que afecten a grandes grupos, así como en investigaciones de carácter cuantitativo. Sin embargo, en el caso de los deficientes la utilidad es menor, ya que no se suele cumplir el requisito básico para que se les pueda aplicar las normas de un proceso de estandarización; es decir, que sean como los individuos a partir de los que se ha elaborado el test. Además, considera que no sirve cuando nos centramos en el individuo, ya que cuando se obtienen bajas puntuaciones puede deberse no a la falta de aptitudes, sino a otras dos razones: no haber tenido oportunidades para aprender las asociaciones y habilidades demandadas por el test, o bien a que los tests no miden adecuadamente las habilidades y estrategias requeridas incluso para el aprendizaje académico.

Independientemente de los postulados ideológicos de cada autor, en lo que existe unanimidad es en considerar que el enfoque tradicional no proporciona información cualitativa útil a la hora de elaborar un programa de entrenamiento individualizado y, por tanto, adaptado a las potencialidades de cada sujeto. Además, el conocimiento del nivel de C.I. puede tener consecuencias negativas para el niño, puesto que puede afectar a las expectativas del maestro, que tenderá a actuar en consonancia con ellas, haciendo que se cumplan, tal y como demostraron Rosenthal y Jacobson (1980). Las repercusiones en el autoconcepto del niño son claras.

2. MODELO DINÁMICO

Se trata de una evaluación, centrada en la modificabilidad de los procesos cognitivos y en la estrecha relación entre el proceso de evaluación e intervención, y hace especial hincapié en detectar en

qué ha fallado el desarrollo cognitivo para proceder a su modificación y conseguir un desarrollo normal. En este sentido es un enfoque mucho más global que el psicométrico tradicional.

Desde una perspectiva histórica, puede considerarse a Vigotsky como el creador de este enfoque. Sin embargo, cuando se analiza la bibliografía al respecto se evidencia una clara disfunción entre el gran desarrollo y profundización que se ha hecho de las teorías elaboradas por Vigotsky y los escasos modelos y pruebas que han sido construidos para evaluar la capacidad cognitiva partiendo del constructo denominado «potencial de aprendizaje». Incluso el propio Vigotsky tampoco profundizó en ese sentido.

Desde nuestro punto de vista, una explicación de ese escaso desarrollo en la construcción de pruebas válidas y fiables a partir de este modelo puede deberse a lo difícil que resulta la validación de los mediadores de aprendizaje que tales pruebas exigen. Igualmente a que este modelo de evaluación diagnóstica requiere, según Brown y French (1979, 261), un detallado análisis de tareas, un apropiado grupo de tareas cognitivas y un análisis de tareas de posibles pruebas de transferencia. Sin esta información será difícil seleccionar series de ayudas graduadas para el aprendizaje original de la tarea, o métodos adecuados para valorar la rapidez y eficacia de la transferencia. Asimismo, se necesita también elaborar baterías referidas a distintos campos, ya que, tal como plantea Vigotsky (1973, 29), el intelecto es más bien la resultante de muchas capacidades diferentes, cada una de las cuales, en cierto modo, es independiente de las otras y por lo tanto ha de ser desarrollada independientemente mediante un ejercicio adecuado.

El diseño para llevar a cabo este tipo de evaluación diagnóstica consiste en: test-entrenamiento-retest. El test corresponde a la medición clásica. El entrenamiento es una práctica de ayuda de escasa duración que es suministrada por el adulto. El retest suele presentar dos modalidades, según se quiera medir el mantenimiento del aprendizaje o la transferencia.

Vigotsky utilizaba el método funcional de la doble estimulación, consistente en plantear al sujeto un problema que no pudiera resolver por sí solo, a tenor de su nivel de desarrollo efectivo, proporcionándole al mismo tiempo medios (fundamentalmente signos) que le ayudaran a resolver el problema. Es decir, el objetivo era comprobar qué podía conseguirse a partir de una determinada intervención. O lo que es lo mismo, pretendía medir la amplitud de la zona de desarrollo potencial.

Campione, Brown, Ferrara y Bryan (1984, 82) intentaron construir un modelo de evaluación a partir de la idea de la zona de desarrollo potencial. Se propusieron comprobar cuánto input se necesitaba para realizar la tarea y cuánto para transferirla a otras. El proceso de evaluación era el siguiente:

Se ponía a trabajar a los sujetos en un problema. Si no lo resolvían, se les daba una serie de indicaciones para ayudarles. Inicialmente eran claves muy generales, que posteriormente se iban especificando y concretando, de tal forma que la última era una clave muy concreta para generar la respuesta correcta. Por último, se estimaba la cantidad de ayuda que habían necesitado.

El mismo proceso emplearon para el estudio de la transferencia.

Haywood (1977), en el laboratorio de Nashville, se ha centrado en el estudio de retrasados mentales ligeros (orgánicos y culturales) en relación a su capacidad de abstracción verbal, entendida como «la actividad de agrupamiento y clasificación de hechos aislados y colocación de etiquetas abstractas en las categorías resultantes». Toma como base que los retrasados mentales puntúan muy bajo en los tests standard, y a la vez hipotetiza que era posible la alteración de las estructuras cognitivas de forma más o menos permanente.

El proceso de evaluación seguido fue el siguiente:

Primero se les pasaba la prueba de semejanzas del test de Wechsler, para obtener la línea base (test).

En segundo lugar, se les pasaba la mitad de los ítems de la prueba, con una presentación enriquecida (coaching), utilizando 3, 4 ó 5 ejemplos.

Finalmente, se les pasaba nuevamente la mitad de los ítems en su forma original, no enriquecida (2 ejemplares: retest).

A partir de este proceso se obtenían 2 índices:

— El grado de modificabilidad: diferencia entre las puntuaciones de la línea base y las obtenidas en la presentación enriquecida.

— El grado de transferencia: diferencia entre la puntuación obtenida en la línea base y la última presentación de ítems en su forma no enriquecida.

En definitiva, lo que midieron fue el potencial de aprendizaje o zona de desarrollo potencial.

Los resultados obtenidos pusieron de manifiesto que en el grupo de retrasados culturales había ganancias cuando se presentaban los ítems en forma enriquecida, pudiéndose deducir, por tanto, que las estructuras cognitivas pueden ser modificadas y, a la vez, que los problemas de abstracción verbal parecen deberse no tanto al procesamiento central sino a la información de entrada que se procesa (deficiencias en el input) en este tipo de niños.

Por otra parte, también comprobaron que el aprendizaje de la fase de presentación enriquecida se transfería a otros ítems en los retrasados mentales culturales, si bien no aparecía en los retrasados mentales orgánicos, bien por insuficiencia del procedimiento o por déficit central, lo cual plantea importantes implicaciones en el campo rehabilitador.

Haywood y Switzky (1974), en un estudio de similares características, habían comprobado que la presentación enriquecida no surtía efectos significativos con los niños aventajados, sino únicamente con los retrasados deprivados culturales, reafirmando, así, la relación entre las exigencias sociales y los procesos de abstracción verbal. Es por ello por lo que los citados autores insisten en la importancia de la mediación del adulto con estas palabras: «No es verdad que la mayoría de los niños aprendan bien si quitáramos los obstáculos mayores y los dejáramos solos; esto no es suficiente; un adulto o un chico mayor debe interpretar para los niños el significado de sus experiencias perceptuales, insistir con prácticas en las que apliquen los principios aprendidos, proveerles de feedback en sus ejecuciones y compartir la emoción de aprender».

De ahí que la evaluación diagnóstica en dicho modelo, al ser una experiencia de aprendizaje, deberá tener especial cuidado en la relación examinador-niño.

Otro autor que ha trabajado extraordinariamente en la elaboración de instrumentos para la evaluación del potencial de aprendizaje es Budoff, quien define la inteligencia (1975,76) como la capacidad para aprovecharse de la experiencia. Esta concepción, como se recordará, es coincidente con la teoría del desarrollo de los procesos cognitivos de Vigotsky (intersubjetividad). Es por ello por lo que algunos niños que han carecido de esa experiencia funcionan como retrasados mentales en los tests tradicionales. Por ello, en su procedimiento de evaluación se brinda esa experiencia a los sujetos, pero controlando los efectos negativos que pueden producirse por la situación del test. En definitiva, de lo que se trata es de trabajar en la zona de desarrollo potencial, pero brindando una ayuda que parta de la «intersubjetividad», que ya hemos comentado anteriormente. entrenamiento. Finalmente, se aplica por tercera vez la prueba, un mes más tarde. La puntuación final obtenida a través de ese procedimiento representaría la capacidad del sujeto para beneficiarse del entrenamiento mediado.

Resulta fácil ver el paralelismo que existe entre los principios del entrenamiento y los conceptos de definición situacional, intersubjetividad y mediación semiótica. Asimismo, puede observarse con claridad la influencia de los conceptos de Vigotsky sobre la importancia del lenguaje como instrumento regulador del pensamiento en el último punto del entrenamiento.

Budoff y Hamilton (1976), cuando trabajaron con deficientes medios y severos, introdujeron una variación en la fase de entrenamiento: entrenamiento dentro del test. Al niño se le proporciona entrenamiento en el momento en que falla un ítem. Se le permiten 3 ensayos, aunque después de cada fracaso se le entrena. Al fallar 6 ítems sucesivos en el primer ensayo se suspende la aplicación.

Sin lugar a dudas, la persona que más ha profundizado en la elaboración de instrumentos apropiados para evaluar el potencial de aprendizaje en niños retrasados mentales ha sido Feuerstein y colaboradores. Por esa razón, nos detendremos más que en los casos anteriores, en la descripción de sus brillantes trabajos.

Este autor parte de una concepción de la inteligencia similar a la de Budoff, aunque hace un mayor hincapié en la importancia del aprendizaje mediado, tal y como se evidencia en esta afirmación: «Cuanto mayor sea la experiencia de aprendizaje mediado y más pronto se someta al niño a dicha experiencia, mayor será la capacidad del organismo para ser modificado y de la misma forma podrá utilizar de forma eficaz toda la estimulación directa» (Feuerstein, 1979). Dicho aprendizaje mediado no es otra cosa que el aprendizaje en la zona de desarrollo potencial, que, tal como ya hemos comentado, va a producir el grado de modificabilidad de las estructuras cognitivas. Por ello, su modelo de evaluación L.P.A.D. (Learning Potential Assessment Device), intenta hacer una valoración de la modificabilidad a través de la introducción de experiencias de aprendizaje mediado que él define textualmente (1979, 71) como: «Los procesos interaccionales entre el organismo humano que se está desarrollando y un adulto con experiencia e intención, quien interponiéndose entre el niño y las fuentes externas de estímulo, le media el mundo, sirviendo de marco, seleccionando, enfocando y realimentando las experiencias ambientales y hábitos de aprendizaje». A la vez, «intenta comprobar la transferibilidad de las estrategias adquiridas a otras áreas de funcionamiento cognitivo distintas a las presentadas en el entrenamiento, e investigar las modalidades preferentes del individuo, con el fin de valorar aquellas en las que el individuo presenta sus puntos más fuertes o más débiles, pudiendo así suplementar unas estrategias preferenciales que conduzcan a la modificación deseada, de la forma más eficiente y directa posible».

La estructura del L.P.A.D. está basada en tres dimensiones: el grado de novedad y complejidad de la tarea, el lenguaje o modalidad de presentación (lógico-verbal, espacial, numérico, pictórico-concreto, figurativo y verbal) y la variedad de operaciones cognitivas que las tareas representan (clasificación, analogía, silogismos, seriaciones, permutaciones...).

El proceso de evaluación es el siguiente: Se presenta al sujeto una tarea y se le da un entrenamiento (aprendizaje mediado) para que la resuelva correctamente. Después se le presentan otras tareas que tienen mayor grado de complejidad, otros lenguajes que requieren otro tipo de operaciones. Así, no sólo se evalúa el grado de modificabilidad, sino también el estilo cognitivo, las áreas de funcionamiento y los lenguajes preferenciales.

Un aspecto de gran importancia para Feuerstein es la situación de examen, que deja de ser la situación fría y neutra entre examinador-examinando, para convertirse en un proceso interactivo adulto-niño, en el que ambos se comprometen en la realización de la tarea, de tal forma que el adulto tiene un papel muy activo ayudando, animando y motivando al niño, pues es la única forma de que se pueda producir la experiencia de aprendizaje mediado.

3. ANÁLISIS DE NUESTRA BATERÍA

Con el fin de no sobrepasar el número de páginas previsto por la comisión organizadora para la presentación de Comunicaciones, no vamos a explicitar el contenido de nuestra Batería, los datos relacionados con su validez, ni tampoco los datos que hacen referencia a su fiabilidad. No obstante, en la exposición será este apartado al que más importancia le concedamos, apoyándonos en una serie de transparencias para facilitar la comprensión del mismo.

4. REFERENCIAS BIBLIOGRÁFICAS

- BROWN, A. L. & FRENCH, L. (1979): «The zone of potential development: Implications for intelligence Testing in the year 2000». *Intelligence*, nº 3, 255-273.
- BUDOFF, M. & HAMILTON, J. L. (1976): «Optimizing test performance of moderately and severely retarded adolescents and adults». *American Journal of Mental Deficiency*, 81, 49-57.
- BUDOFF, M. (1975): «Measuring learning potential: an alternative to the traditional intelligence test». En GRED, G. (ed.): *Ethical and legal factors in the practice of School Psychology. Proceeding of the first annual conference in School Psychology*.
- CAMPIONE, BROWN, FERRARA & BRYANT (1984): «The zone of proximal development: implications for individual differences and learning». En ROGOFF y WERSTCH (eds.): *Children's Learning in the «Zone of Proximal Development»*. *New directions for child development*, San Francisco, Jossey Bass, nº 23.
- DAS, J. P., KIRBY, J. & JARMAN, R. F. (1979): *Simultaneous and succesive cognitive processes*. New York, Academic Press.
- FEUERSTEIN, R., RAND, Y. & HOFFMAN, M. D. (1979): *The dynamic assessment of retarded performers: the learning potential assessment device, theory, instruments and techniques*. University Park Press, Baltimore.
- HAYWOOD, H. C. & SWITZKY, H. N. (1974): «Children's Verbal Abstracting: Effects of Enriched Input, Age and IQ». *American Journal of Mental Deficiency*, 78 (5), 556-565.
- HAYWOOD, C. (1977, a): «A cognitive approach to the education of retarded children». *Peabody Journal of Education*, 54, 110-116.
- HAYWOOD, C. (1977, b): «Alternatives to normative assessment». *Research to Practice in Mental Retardation*. Peter Mittler.
- MISSIUNA, CH. & SAMUELS, M. (1988): «Dynamic Assessment: Review and Critique», *Special Services in the Schools*, Vol. 5 (1/2), 1-21.
- MOLINA, S. & ARRAIZ A. (1993, a): *Procesos y estrategias cognitivas en el niño deficiente mental*. Pirámide, Madrid.
- MOLINA, S., ARRÁIZ, A. & GARRIDO, M. A. (1993, b): *Batería para la Evaluación Dinámica del Potencial de Aprendizaje y de Estrategias Cognitivas (BEDPAEC)*.
- VIGOTSKY, L. S. (1973): «Aprendizaje y desarrollo intelectual en la edad escolar». En LURIA, A. R. y LEONTIEV, A. N. (eds.): *Psicología y Pedagogía*, 23-39. Akal, Madrid.

PROCESO DE CONSTRUCCIÓN Y VALIDACIÓN DEL TEF

Cristina Cardona Moltó, M. Ángeles Martínez Ruiz, Narciso Sauleda Parés
Universidad de Alicante

INTRODUCCIÓN

La valoración de las dificultades escolares continúa siendo, hoy día, cuestión crítica para el logro de una respuesta adecuada a las necesidades educativas que los alumnos plantean.

Habitualmente, el proceso de valoración incluye información obtenida a través de tests estandarizados de inteligencia y rendimiento, así como, dibujos y análisis de tareas. Sin embargo, autores como Keogh, Becker, Kukic y Kukic (1974) advirtieron de la limitada utilidad que proporciona dicha información para orientar los procesos de enseñanza-aprendizaje y el trabajo real con el alumno en el aula.

Asimismo, la introducción del término 'necesidades educativas especiales' aplicable a una mayor proporción de niños (20% de la población escolar) contribuye a que el propósito y los métodos de valoración se hallen más intrínsecamente relacionados con la planificación de la intervención que con la categorización de las dificultades.

Por tanto, mientras el nivel de habilidades adquiridas representa un cúmulo importante de variables a considerar resulta, asimismo, de interés el examen de las características funcionales del niño. A este respecto, la investigación ha demostrado que los niños con problemas de aprendizaje y conducta difieren de aquéllos otros sin problemas, no sólo en lo que a habilidades adquiridas se refiere, sino también en el **modo** de realizar las tareas escolares, actividad que desarrollan en función de ciertas características individuales como impulsividad, problemas atencionales, patrones de comportamiento temperamentales y problemas motivacionales que pueden exacerbar sus problemas en el contexto escolar (Palmer, 1980).

En relación a la impulsividad, algunos trabajos vienen a demostrar que los sujetos reflexivos reúnen la información de un modo más cuidadoso y sistemático que los impulsivos y que su relación con el rendimiento es evidente como lo demuestra el hecho de una mayor impulsividad en los niños con dificultades lectoras (Kagan, 1965), problemas de aprendizaje (Keogh y Donlon, 1972) y fracaso escolar (Messer, 1970).

Epstein, Hallahan y Kaufman (1975), en su revisión de la literatura sobre implicaciones de la dimensión reflexividad-impulsividad para la Educación Especial, presentan evidencias que sugieren que mientras el bajo rendimiento académico de los niños con dificultades ha sido frecuentemente atribuido a una falta de capacidad o conocimiento, la investigación por ellos revisada apunta que es un **tempo** cognitivo impulsivo el factor que puede explicar, en parte, el pobre aprendizaje de algunos sujetos con problemas.

Sin embargo, los resultados de otros estudios muestran que los niños con discapacidad de aprendizaje y normales no difieren en el **tempo** cognitivo, tal como ha sido operativizado en el **Matching Familiar Figures Test**, pero si lo hacen en la proporción de errores en el mismo test. Estos resultados, en concordancia con los hallados por Haskins y McKinney (1976), sugieren que el parámetro '**exactitud de respuesta**' es un mejor predictor del rendimiento que el promedio de latencia, indicando que los niños con dificultades de aprendizaje no son más impulsivos, sino que más bien hacen un uso pobre de las estrategias de procesamiento de la información, estrategias que, por otro lado, pueden ser modificadas (Messer, 1976).

Para clarificar estas cuestiones y ante la capacidad moderadora de esta variable, estimamos de gran interés el estudio del constructo reflexividad-impulsividad, teniendo en cuenta que el análisis de esta dimensión cognitiva nos pone en condiciones de aportar una orientación más enriquecedora al diagnóstico y la valoración de los problemas de aprendizaje escolar, así como, nos brinda la posibilidad de mejorar la comprensión de los procesos de enseñanza-aprendizaje.

Así pues, nuestra aportación consiste en valorar críticamente el TEF (Test de Emparejamiento de Figuras) como instrumento, elaborado por nosotros, con el intento de superar los errores metodológicos asociados al «Matching Familiar Figures Test» MFFT (Kagan, Rosman, Day, Albert y Phillips, 1964), prueba originariamente diseñada para evaluar este constructo.

Elaboración del TEF

El Test de Emparejamiento de Figuras constituye una versión propia (Cardona, 1993) del MFFT. Este es un test de emparejamiento perceptivo que se realiza a través de tareas de discriminación visual. Consta de 15 ítems a través de los cuales se le presenta al sujeto el dibujo de una figura que le resulta familiar (casa, reloj, TV,...) y seis variantes de la misma de las que sólo una es la copia del modelo estándar. La tarea consiste en identificar la variante que es idéntica al modelo.

Se toman dos medidas de la ejecución del sujeto: 1) El tiempo que tarda en dar su primera respuesta (latencia) y 2) El nº de errores que comete en cada ítem hasta que acierta o hace un total de seis elecciones incorrectas.

Seguidamente, se obtiene el promedio de ambas variables para los 15 ítems que componen la prueba. El tiempo de aplicación suele ser de unos 15 minutos aproximadamente.

Típicamente, el sujeto que, consistentemente, emplea más tiempo en responder y comete menos errores (por encima de la mediana del grupo en tiempo y por debajo en errores) se define como reflexivo, mientras que, el que, regularmente, responde de manera rápida e inexacta (por debajo de la mediana en tiempo y por encima en errores) es clasificado como impulsivo.

Esta prueba fue diseñada para un estudio previo, dadas las limitadas garantías técnicas del test original, el MFFT, con coeficientes test-retest que oscilaban entre .23 (Messer, 1970) y .43 (Siegelman, 1969) sobre unos intervalos de tiempo de entre 3 semanas y 2 1/2 años; y unos coeficientes alfa de consistencia interna situados entre .32 y .60 para la puntuación de error, en contraste con los encontrados para la latencia (.89) (Ault, Mitchell y Hartmann, 1976).

Otros problemas derivan de la moderada correlación negativa entre latencia y errores (a más latencia menos cantidad de errores) con un promedio en las diversas investigaciones citadas por Ault y col. (1976) de algo más de $-.50$.

Ante estos inconvenientes optamos por elaborar una versión más fiable incorporando a nuestra prueba un número mayor de ítems y aumentando el poder discriminativo de los mismos a través del diseño de dibujos predominantemente geométricos y no tanto artísticos. Para el cálculo del índice de discriminación se procedió a la obtención de la correlación de cada ítem con la puntuación total del test

mediante el coeficiente *r* biserial puntual. El análisis de los ítems dio origen a la eliminación de tres de ellos y a una nueva reordenación de los mismos.

Los índices de fiabilidad como consistencia interna obtenidos en una aplicación previa a 60 escolares de 6-9 años con necesidades especiales para el aprendizaje de las técnicas instrumentales fueron: .85 para la latencia y .77 para los errores; la correlación entre errores y latencia se vio, asimismo, sensiblemente mejorada (-.65).

Validación del TEF

Para el proceso de validación existe, en curso, una investigación (Cardona, Martínez y Sauleda, 1993) que persigue los siguientes objetivos:

- 1) Examinar cuál es la relación existente entre la dimensión reflexividad-impulsividad medida a través del TEF y el rendimiento académico en 1º y 2º cursos de Primaria (validez concurrente).
- 2) Calcular la fiabilidad test-retest del TEF.
- 3) Averiguar su poder predictivo.
- 4) Valorar la relación existente entre las variables exactitud de respuesta y latencia.

En estrecha relación con los objetivos se formularon cuatro hipótesis:

- H1)** Existe relación entre el rendimiento en lectura, escritura y cálculo y el nº de errores en el TEF, pero no la habrá entre rendimiento y latencia. Por consiguiente, los alumnos más exactos pondrán de manifiesto un rendimiento significativamente superior al que exhiban los alumnos que cometen más errores en la ejecución del TEF.
- H2)** Existe relación significativa entre los alumnos clasificados como impulsivos o reflexivos en 1º y 2º cursos de Enseñanza Primaria (fiabilidad test-retest),
- H3)** Existe una relación significativa entre reflexividad-impulsividad (nº de errores en el TEF, no latencia) y rendimiento académico en 1º y en 2º de Primaria.
- H4)** Existe una correlación moderada-alta entre errores y latencia que se espera favorecerá más a los niños que a las niñas y a los escolares de 2º que a los de 1º.

La muestra está formada por la totalidad de la población escolar de **primer curso de Primaria** procedente de los **tres únicos colegios** de una localidad alicantina, población de unos 13.000 habitantes situada a 54 Km. de Alicante. Estimamos estará integrada por 170-180 alumnos (edades comprendidas entre 6-8 años, Primer Ciclo de Primaria) procedentes de los seis cursos de 1º y los respectivos 12 maestros de nivel.

Todos los datos serán recogidos de la misma muestra al concluir 1º y 2º de Primaria, por tanto, el proyecto de validación abarca dos cursos escolares.

Las variables sometidas a estudio son: la **conducta reflexivo-impulsiva**, que se evaluará a través del TEF y el **rendimiento** en lectura, escritura y cálculo, indicadores, estos últimos, que se evaluarán a través de dos técnicas: 1) Mediante una escala de evaluación que cumplimentará el maestro y 2) A través de un conjunto de pruebas (algunas estandarizadas) que se aplicarán a los alumnos.

Por tanto, al concluir 1º se administrarán las pruebas:

- **TALE.** Test de Análisis de la Lectoescritura (Toro y Cervera, 1984). Subtests de Lectura: Letras, sílabas palabras y texto 1A y 1B; y subtest comprensión lectora (Texto I).
- **PEP.** Prueba de escritura de palabras simples (Cardona, 1993) y Dictado 1 del TALE.
- Prueba **OPER** de operaciones y problemas (Cardona, 1993).

Este conjunto de pruebas nos permitirá obtener un índice de la validez concurrente a través de la puntuación obtenida con las escalas de evaluación (TINS) construidas 'ad hoc'.

Al término de 2º curso de Primaria se valorará a los mismos alumnos con los siguientes instrumentos:

- **TALE.** Subtets de Lectura: Letras, sílabas, palabras y texto 2; y subtest comprensión lectora (Texto II).
- Subtest Alteraciones de la Escritura (A.E) de la Batería BADYG-B (Yuste, 1984).
- Subtest Aptitud para el Cálculo (Ap.N) de la Batería BADYG-B.

Los datos serán analizados en cuatro fases. Los análisis efectuados en la **Fase I** nos conducirán al contraste de la hipótesis primera, empleando la técnica de correlación producto-momento de Pearson. Se correlacionará:

- La puntuación (errores y latencia) en el TEF con las puntuaciones obtenidas de la Escala de Evaluación TINS en 1º.
- La puntuación (errores y latencia) en el TEF con las obtenidas a través de las pruebas TALE y BADYG-B en 2º curso.

En la **Fase II** se empleará igualmente la técnica de correlación de Pearson para el cálculo de la estabilidad temporal del TEF. Las variables serán la latencia y errores (total grupo, niños y niñas) en los pases efectuados en 1º y 2º. El tiempo entre aplicaciones será de un año.

La fiabilidad test-retest será, asimismo, examinada mediante un análisis de contingencia para determinar el grado en que los alumnos clasificados como reflexivos o impulsivos en 1º continúan siéndolo en 2º curso.

En una **tercera fase** se procederá al cálculo de la validez predictiva correlacionando la puntuación de error y latencia en el TEF obtenidas en primer curso, con las puntuaciones en rendimiento (lectura, escritura y cálculo) obtenidas en 2º. De este modo, podremos comprobar en qué grado el TEF nos sirve como instrumento para predecir las dificultades de aprendizaje.

En último lugar, se calcularán las correlaciones (1ª y 2ª administración del TEF) entre los errores y la latencia para el total del grupo, niños y niñas, en respuesta al cuarto contraste de hipótesis.

REFERENCIAS BIBLIOGRÁFICAS

- AULT, R. L., MITCHELL, C. y HARTMANN, D. P. (1976): Some methodological problems in reflection-impulsivity research. *Child Development*, 47, 227-231.
- CARDONA, M. C. (1993): *Aulas de apoyo e integración escolar. Evaluación de un programa alternativo de apoyo para alumnos con necesidades especiales*. Facultad de Filosofía y CC.EE. Dpto. MIDE. UNED: Madrid. Tesis Doctoral inédita.
- CARDONA, M. C., MARTÍNEZ, M. A. y SAULEDA, N. (1993): *Reflexividad-impulsividad y rendimiento académico: Aportaciones del TEF a la detección y valoración de las dificultades escolares*. Trabajo presentado al Concurso Nacional de Proyectos de Investigación Educativa de 1993. CIDE. Madrid.
- EPSTEIN, M. H., HALLAHANN, D.P. y KAUFFMAN, J. M. (1975): Implications of the reflectivity-impulsivity dimension for special education. *The Journal of Special Education*, 9 (1), 11-25.
- HASKINS, R. y MCKINNEY, J. (1976): Relative effects of response tempo and accuracy on problem-solving and achievement. *Child Development*, 47, 690-696.
- KAGAN, J. (1965): Reflection-impulsivity and reading ability in primary grade children. *Child Development*, 36, 609-628.
- KAGAN, J., ROSMAN, B. L., DAY, D. ALBERT, J. y PHILLIPS, W. (1964): Information processing in the child: Significance of analytic and reflective attitudes. *Psychological Monographs*, 78 (1, whole n. 578).

- KEOGH, B. K., BECKER, L. D., KUKIC, M. y KUKIC, S. (1974): Programs for EH and EMR pupils: Review and recommendations, Part I. *Academic Therapy*, 1 (3), 176-198.
- KEOGH, B. y DONLON, G. (1972): Field dependence, impulsivity and learning disabilities. *Journal of Learning Disabilities*, 5, 331-336.
- MARTÍNEZ RUIZ, M. A. (1990): Análisis y taxonomización de actividades discentes. *Enseñanza. Anuario Interuniversitario de Didáctica*, 8, 47-56.
- MESSER, S. B. (1970): Reflection-impulsivity: Stability and school failure. *Journal of Educational Psychology*, 61, 487-490.
- MESSER, S. B. (1976): Reflection-impulsivity: A review. *Psychological Bulletin*, 83 (6), 1.026-1.052.
- PALMER, D. J. (1980): Factors to be considered in placing handicapped children in regular education classes. *Journal of School Psychology*, 18 (2), 163-171.
- SIEGELMAN, E. (1969): Reflective and impulsive observing behavior. *Child Development*, 40, 1.213-1.222.
- TORO, J. y CERVERA, M. (1984): *TALE. Test de análisis de lectoescritura*. Madrid: Visor Libros.
- YUSTE, C. (1984): *Manual técnico BADYG-B*. Madrid: CEPE.

DIAGNÓSTICO DE LAS DIFERENCIAS ÉTNICAS Y DE LOS PROCESOS DESARROLLADOS EN LA EDUCACIÓN PRIMARIA. PERCEPCIÓN DEL PROFESORADO

Flor Cabrera, Julia V. Espín, M^a. Ángeles Marín y Mercedes Rodríguez Lajo
Universidad de Barcelona

1. INTRODUCCIÓN

La finalidad de esta comunicación es presentar la Guía de Entrevista para conocer la *percepción de profesorado* sobre su experiencia en un contexto multicultural. Se trata de uno de los instrumentos evaluativos que se han elaborado para llevar a cabo un estudio sobre «*El diagnóstico de la población infantil étnicamente diferenciada*» en Barcelona. Este estudio forma parte de una investigación más amplia que se sitúa en el marco de la educación intercultural: «*Diagnóstico de las diferencias étnicas y de los procesos desarrollados en la educación primaria*»¹ investigación que comprende dos grandes objetivos:

A) Realizar un diagnóstico de la población infantil inmigrante étnicamente diferenciada, escolarizada en la educación primaria en Barcelona.

B) Comprender en profundidad las relaciones educativas que se establecen en las escuelas donde se inserta la población estudiada mediante la metodología etnográfica.

El primer objetivo se cubrirá mediante una *investigación por encuesta* que abarca el estudio de:

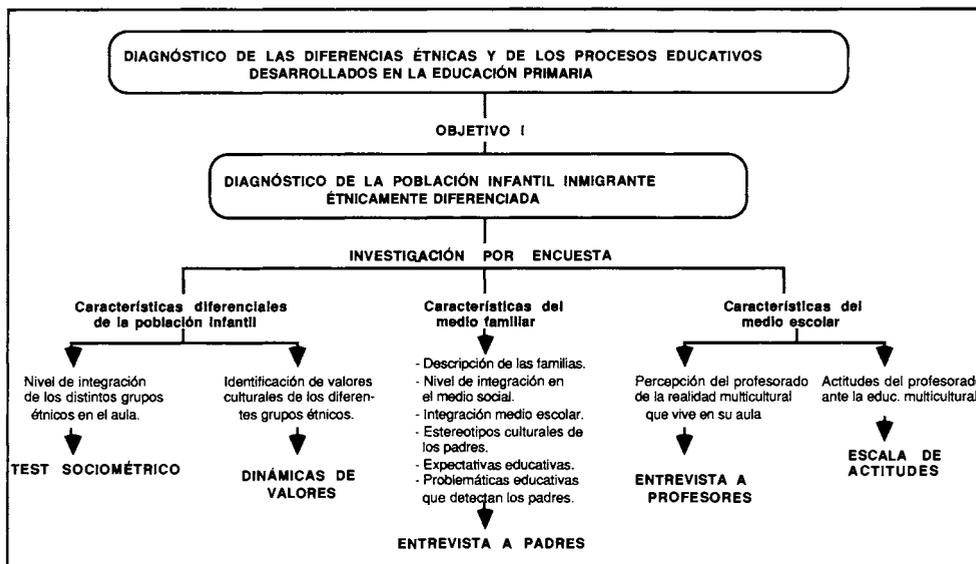
- 1.— Las características diferenciales de la población infantil de las etnias seleccionadas.
- 2.— Las características del medio escolar.
- 3.— Las características del medio familiar.

En el cuadro que se incluye a continuación, se recoge un resumen del contenido de los instrumentos utilizados en esta investigación. En las páginas siguientes y en otras dos comunicaciones se detallan con amplitud las fases seguidas para la elaboración de los instrumentos utilizados para el diagnóstico del medio escolar y del medio familiar.

2. IDENTIFICACIÓN DE LA POBLACIÓN INMIGRANTE ÉTNICAMENTE DIFERENCIADA

No es fácil obtener unos datos estadísticos donde se muestre la realidad de la inmigración en Cataluña. Las fuentes de información han sido:

1 Investigación subvencionada por el C.I.D.E. y que está siendo realizada por el siguiente equipo de investigación: Margarita Bartolomé —directora del proyecto—, Flor Cabrera, Julia Victoria Espín, María Ángeles Marín, Juana Martín, Delio del Rincón y Mercedes Rodríguez Lajo. Profesores en su mayoría del Dpto. M.I.D.E. de la Universidad de Barcelona.



- Delegación Provincial de Trabajo y Seguridad Social de Barcelona. Memoria 1991.
- La última Regulación de inmigrantes en España. Barcelona.
- El Padrón Municipal de 1990.
- El Censo realizado en las escuelas de los niños escolarizados de otras nacionalidades, realizado por el Departament de Ensenyament de la Generalitat. Datos de Barcelona ciudad.

En la selección de la procedencia de la población inmigrada se ha tenido en cuenta la distribución que realiza el Padrón Municipal y la última regularización. Esta población se ha agrupado en tres grandes bloques:

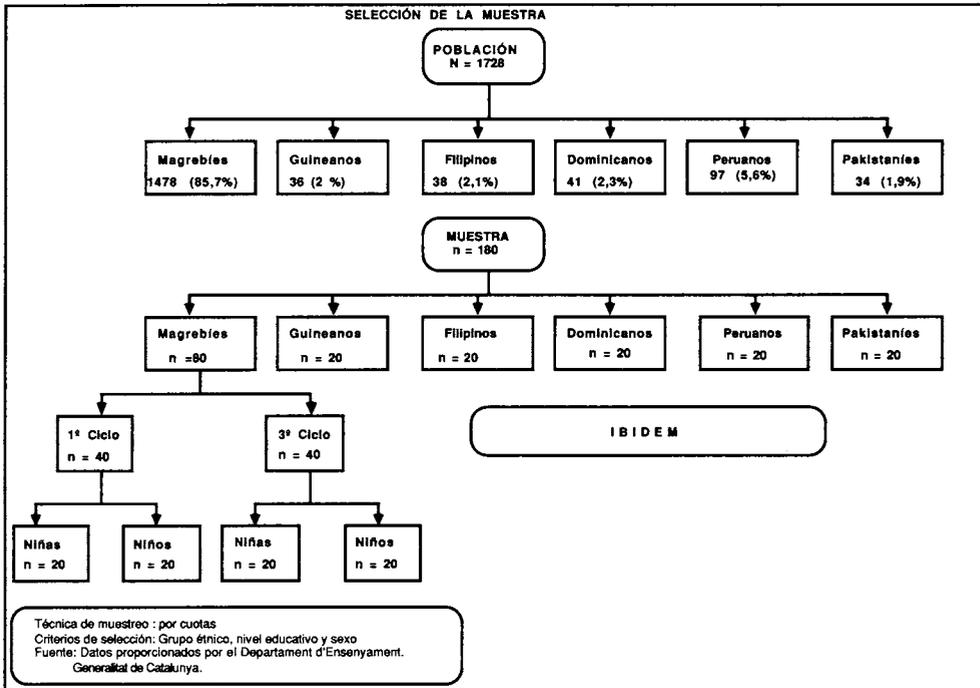
- América (Países de América latina)
- África (Subsahariana y el Magreb)
- Asia (Extremo Oriente).

Dentro de estos tres grandes bloques se seleccionan aquellos países de los que hay un mayor número de inmigrantes. Si tenemos en cuenta las variables número de población y país de procedencia, el gráfico es el siguiente:

GRUPOS DE LA POBLACIÓN ELEGIDOS

Pais	Población escolar
Perú	164
República Dominicana	65
Marruecos	1560
Guinea Ecuatorial	42
Filipinas	69
Paquistán	38

La población total infantil escolarizada en centros, con edades comprendidas entre 6 y 12 años es de 1.728 niños. De esta población se ha escogido una muestra de 180 niños como objeto de investigación. Siguiendo la técnica de muestreo por cuotas, y aplicando los criterios de grupo étnico, nivel educativo y sexo, la distribución queda según muestra el cuadro siguiente.



3. FINALIDAD DEL ANÁLISIS DE LA PERCEPCIÓN DE PROFESORADO

Nuestro objetivo general es conocer en qué medida el profesorado que trabaja en aulas con alumnos de diferentes etnias realiza un planteamiento curricular y didáctico que contemple la diversidad multicultural. En concreto, tratamos de analizar cuál es la naturaleza de sus planteamientos pedagógicos diferenciales, qué valoración hace de los mismos, qué medios y recursos utiliza para ello y cómo valora la efectividad de estos medios y recursos.

Distinguiremos dos momentos distintos de la planificación docente que puede dar lugar a problemáticas también distintas en el objetivo de nuestro estudio, a saber:

a) cómo el profesor afronta la programación inicial de las actividades del aula cuando tiene alumnos de diferentes etnias.

b) cómo afronta el transcurrir del curso con niños de diferentes etnias en la clase.

Tomamos en consideración como objetivo importante también de nuestro estudio las diferencias que puede haber entre los profesores en razón a aspectos tales como: experiencia como profesor en situaciones multiculturales, experiencia del centro con aulas multiculturales, tipología del centro. Así como las diferencias que pudieran resultar según los tipos de etnias presentes en el aula.

Los objetivos generales que configurarían el estudio podemos ubicarlos en los grandes ámbitos siguientes.

1.— Proceso de adscripción de las diferentes etnias a las aulas.

Bajo este epígrafe tratamos de conocer en qué medida los profesores están de acuerdo con los criterios que se utilizan en los centros para adscribir a los niños a una u otra aula, y en caso negativo identificar cuáles deben ser éstos, específicamente:

- Naturaleza de los criterios utilizados.
- Idoneidad de los criterios.

2— Planificación inicial de las actividades escolares.

Se trata de conocer aspectos esenciales que en una planificación de curso con aulas multiculturales deben tomarse en consideración:

— En qué medida los profesores disponen de un diagnóstico inicial de los alumnos de las diferentes etnias en los aspectos que consideran relevantes para programar las actividades escolares de estos alumnos.

— En qué medida las necesidades de diagnóstico inicial que expresan los profesores responde a un planteamiento docente multicultural.

— Apoyos específicos que tienen para llevar a cabo la programación inicial y cómo valoran tales ayudas.

— En qué medida en el planteamiento inicial que hace el profesor de la organización del aula atiende a las necesidades de las minorías étnicas, así como los criterios de distribución de los alumnos en el aula y su estabilidad a lo largo del curso, distinguiendo tipos de agrupamientos según los criterios los imponga el profesor o los propios alumnos, como indicadores de planteamientos multiculturales en el aula.

Estos objetivos suponen elaborar preguntas que guíen la entrevista respecto a:

— Nivel de conocimientos que tenía el profesor relativos a: aspectos personales, familiares y de la cultura de origen de sus alumnos.

— Existencia de programación inicial diferencial para los alumnos de las minorías étnicas.

— Ayudas específicas recibidas para la programación inicial, nivel de satisfacción y deficiencias de las mismas.

— Cambios introducidos los primeros días de clase por tener presentes minorías étnicas en el aula.

— Distribución de los alumnos en el aula, criterios de distribución y dinámica de los mismos a lo largo del curso.

— Diferencias en la programación inicial para las diferentes etnias según el dominio que tengan del lenguaje instrumental.

3 — Naturaleza de la intervención individualizada que presta el profesor a los niños de las minorías étnicas y grado en que hace un planteamiento educativo multicultural a nivel de aula durante el curso

Se trata en este punto de conocer en qué medida el profesor realiza una integración de los niños de las minorías étnicas mediante una adaptación de su intervención a las características de estos niños y no se queda en un mero «rebajar los niveles», o bien en qué medida trasciende a la integración haciendo un planteamiento intercultural en su aula.

La concreción de estos objetivos en la guía de entrevista supone plantear cuestiones relativas a:

— La naturaleza de las actividades que hacen los alumnos de las diferentes etnias presentes en el aula, distinguiendo si tienen un dominio suficiente o no de la lengua instrumental.

— Tratamiento curricular que se le debe dar al «conocimiento de las diferentes culturas».

— Forma de afrontar los conflictos que pueden darse entre los alumnos, grado en que los utiliza para el desarrollo de la comprensión y respeto entre las diferentes culturas.

— Existencia de pautas de conductas diferenciales que atiendan necesidades específicas de las diferentes culturas.

— Grado en qué la experiencia en aulas multiculturales ha llevado al profesor a replantearse su forma habitual de enseñanza, y tipo de innovaciones pedagógicas introducidas.

— Naturaleza de las ayudas y asesorías específicas utilizadas durante el curso, diferenciando según el nivel lingüístico de los alumnos de las minorías étnicas, distinguiendo entre internas (del propio centro escolar) y externas (fuera del centro).

— Valoración de la efectividad de tales ayudas y asesorías.

4.— Valoración de los recursos humanos y materiales específicos para situaciones educativas multiculturales.

En este punto pretendemos conocer el uso que el profesorado hace de los recursos humanos y materiales específicamente elaborados para contextos multiculturales y que actualmente están a su disposición, y cómo valora su efectividad pedagógica. Es interesante, también, determinar en qué medida el profesorado identifica como principal problema el lingüístico, dejando como secundario las ayudas que pudiera recibir de los especialistas para hacer en sus aula planteamientos multiculturales realmente comprensivos de la diversidad.

La concreción de estos objetivos en términos de preguntas que guíen la entrevista hacen referencia a:

— Recursos existentes a nivel de centro para afrontar un trabajo de aula multicultural.

— Grado de efectividad de tales recursos.

— Grado en que ha utilizado recursos externos al centro.

— Grado en que valora la efectividad de los recursos externos al centro que ha utilizado.

— Situaciones de aula en las que considera que la ayuda de un especialista es totalmente necesaria.

5.— La relación profesores-padres.

En este epígrafe tratamos de conocer la naturaleza de la relación que mantiene el profesor con los padres, la evolución que durante el curso puede haber tenido y en qué medida existen diferencias entre las etnias.

Estos objetivos quedan concretados en la guía de entrevista mediante preguntas relativas a:

— Frecuencia de relaciones

— Formas de establecer la relación

— Temas de conversación

— Dificultades de la relación

— Diferencias entre etnias

— Evolución de la relación.

6.— Valoración general de la experiencia como profesor/a de un aula multicultural.

Tratamos en este punto de conocer la valoraciones sobre limitaciones y ventajas que pueden darse en un aula multicultural y la posible evolución de su importancia a lo largo del curso.

Para estos objetivos las preguntas que guían la entrevista hacen referencia a:

— Grado en que el profesorado ha vivido, distinguiendo entre el inicio y durante el curso, una serie de situaciones que los estudios señalan como representativas de ventajas o de limitaciones en aula multicultural.

— Aspectos que destacan como facilitadores de la integración de los niños de las minorías étnicas en la dinámica escolar. Diferenciando entre etnias si lo considera conveniente.

— Aspectos que destacan por dificultar o limitar la integración de estos niños en la dinámica escolar. Diferenciando entre etnias si lo considera conveniente.

— Nivel de satisfacción global de su experiencia de trabajo en aulas multiculturales.

7.— Percepción del profesor sobre los atributos que caracterizan a cada etnia.

Tratamos de conocer en qué medida el profesor identifica rasgos personales destacables de las diferentes minorías étnicas consideradas como grupo.

Este objetivo se concreta en una pregunta guía donde el profesorado debe señalar los atributos que destacan de las minorías étnicas con las que ha trabajado en relación a:

- Capacidades cognitivas
- Rasgos personales
- Hábitos
- Habilidades sociales

4. METODOLOGÍA

La recogida de información se realiza mediante entrevista personal con el profesorado. Se trata de una entrevista semiestructurada en cuanto que el profesorado dispondrá de un guión de entrevista paralelo al que tiene el entrevistador, distinguiendo así:

a) El guión de la entrevista para el profesorado: donde se plantean las preguntas prácticamente todas ellas de forma abierta a fin de no orientar sus respuestas con las categorías establecidas.

b) El guión de entrevista para el entrevistador: donde las mismas preguntas se presentan con formato cerrado en categorías de respuestas —siempre dejando la oportunidad de que se den en el transcurso de la entrevista nuevas categorías— a fin de facilitarle el desarrollo de la entrevista y una rápida codificación de las respuestas dadas por los profesores.

4.1. Proceso de elaboración de las guías de entrevistas

Una vez delimitado los ámbitos sobre los que se recogerían la información, y especificados los objetivos que interesaban, se elaboró una primera guía de entrevista bastante exhaustiva. Dicha guía se aplicó a 4 profesores, especialmente escogidos por su experiencia como profesores de aulas multiculturales e interesados en el tema. Con esta aplicación piloto se pretendía que se valorara desde la experiencia práctica de los profesores:

- relevancia del contenido tratado en la guía, su secuenciación y estructuración.
- formato de las preguntas.
- idoneidad de la longitud y criterios a seguir para acortarlo en caso necesario.
- preguntas que deberían quedar de forma abierta para constituir el guión de entrevista del profesor.

En función de los resultados del estudio piloto, se realizaron las modificaciones oportunas, reelaborando definitivamente la guía de entrevista para el entrevistador y la guía de la entrevista para el profesor.

BIBLIOGRAFÍA

- AGUADO, M. T. (1991): La educación intercultural: concepto, paradigmas, realizaciones en C. JIMÉNEZ (Coor.): *Lecturas de Pedagogía Diferencial* (pp. 89-104). Madrid: Dykinson.
- ALEGRET, J. L. (1992): Racismo y educación en P. FERMOSO (Ed.): *Educación intercultural: La Europa sin fronteras* (pp. 93-111). Madrid: Narcea.
- BANKS, J. A. y BANKS, C. A. (1989): *Multicultural Education. Issues and perspectives*. Boston: Allyn and Bacon.
- BARTOLOMÉ, M. (1992): Diseños y metodología de investigación desde la perspectiva de la educa-

- ción intercultural. X Congreso Nacional de Pedagogía. *Educación intercultural en la perspectiva de la Europa Unida*. Tomo II. (pp. 647-674). Salamanca: Imprenta provincial.
- CARTER, R. T. (1991): Cultural values: A review of empirical research and implications for counseling. *Journal of Counseling and Development*, 70, 1, 164-173.
- D'ANDREA, M. DANIELS, J. y HECK, R. (1991): Evaluating the impact of Multicultural Counseling Training. *Journal of Counseling and Development*, 70, 1, 143-150.
- EDUCACIÓ COMPENSATÒRIA (1992): *L'escolarització dels nens i nenes àrabs. Orientacions*. Barcelona: Departament d'Ensenyament. Generalitat de Catalunya.
- LOPERA, C. (1991-92): *El bagatge cultural dels magribins*. Programa d'Educació Compensatòria. Barcelona: Departament d'Ensenyament. Generalitat de Catalunya.
- MORERES, J. y SERRA, C. (1992): *El tractament de la diversitat cultural en el material escolar*. Programa de Educació Compensatòria. Barcelona: Departament d'Ensenyament. Generalitat de Catalunya.
- PASCUAL i SAUC, J. (1993): *La immigració estrangera a Catalunya. L'Estat de la Qüestió*. Barcelona: Fundació Jaume Bofill.
- TAYLOR, S. L. y BOGDAN, R. (1986): *Métodos de investigación educativa*. Madrid: La Muralla.

DIAGNÓSTICO DEL CONTEXTO FAMILIAR DE ALUMNOS DE MINORÍAS ÉTNICAS¹

Flor Cabrera, Julia V. Espín, M^a Ángeles Marín y Mercedes Rodríguez Lajo
Universidad de Barcelona

1. FINALIDAD

Pretendemos caracterizar el medio familiar de las etnias seleccionadas, para establecer una descripción y comparación, inter e intra étnica con relación a una serie de variables familiares que hacen referencia a su nivel de integración en el medio social en el que viven, su relación con la cultura de origen, el mantenimiento de costumbres y valores propios, así como sus actitudes hacia la escuela o la problemática que detectan en ella respecto a sus hijos.

Nos interesa comprobar en que medida hay o no, diferencias en cuanto a las pautas culturales y educativas familiares entre las diferentes etnias, y por otra parte queremos detectar su percepción e integración en el medio escolar.

El estudio del contexto familiar, de las pautas educativas de los padres, de la diferenciación de roles según el sexo, y de la problemática percibida por ellos en la escuela se contrastará con los datos aportados por los profesores de estos niños y con la situación de estos últimos dentro del aula y su rendimiento. A partir de los resultados obtenidos, tanto en el diagnóstico del medio familiar como escolar esperamos poder ofrecer algunas pautas de intervención, tanto a las familias como a los centros para favorecer la educación multicultural.

Los objetivos generales y específicos de nuestra investigación en relación con las familias son los siguientes:

1.1. Objetivos generales

1. Conocer las características del medio familiar. En este apartado nos interesa sobre todo obtener información del contexto familiar al que responde cada una de las etnias seleccionadas. Caracterizaremos a la muestra en función de variables como la edad, tipología familiar, lugar de nacimiento, número de hijos, residencia y escolaridad de los mismos, nivel de formación y situación laboral de los padres, sin olvidar la situación legal y su proyecto de migración.

¹ El trabajo que se presenta en esta comunicación se enmarca dentro de una investigación más amplia que lleva por título *Diagnóstico de las diferencias étnicas y de los procesos desarrollados en la educación primaria*, por ello la bibliografía queda recogida en la comunicación que hace referencia a la percepción del profesorado.

2. Nivel de integración de la familia en el medio social próximo. Tipo de relaciones que establecen con las personas de la propia cultura residentes aquí y con las personas del país de acogida. Valoración y mantenimiento de las propias costumbres e introducción de costumbres propias del país de acogida.

3. Identificar los valores y estereotipos culturales de los padres. Pretendemos profundizar en el conocimiento de los aspectos diferenciales de cada etnia en lo referente a la diferenciación sexual de papeles, roles de autoridad y toma de decisiones y la percepción de estereotipos referentes a características personales.

4. Conocer las expectativas educativas respecto de los hijos. En este punto nos interesa conocer el valor que atribuyen a la escuela en la formación de los hijos, es decir cuál creen que es según ellos la función que debe desempeñar la escuela y sobre qué debe centrarse la educación en ella. También nos interesa conocer su percepción con respecto al nivel de formación que ven posible de alcanzar por sus hijos, y las expectativas respecto a su cualificación profesional futura.

5. Identificar el nivel de integración y actitudes de los padres hacia la escuela. Saber qué conocimiento tienen tanto del sistema educativo como de la escuela en particular y su nivel de participación en ella. Nos interesa saber sus actitudes hacia el tipo de educación que reciben los hijos y el modo de tratar la propia cultura en la escuela.

6. Problemáticas educativas de los hijos detectadas por los padres. Los aspectos que nos interesan están relacionados con las dificultades percibidas por los padres respecto de la escolaridad de los hijos, tanto en lo que se refiere a los niveles académicos alcanzados como a la socialización. También nos parece importante conocer el grado de satisfacción de las relaciones de los padres con el centro.

2. METODOLOGÍA

Se acordó que la recogida de la información se efectuaría mediante una entrevista personal con cada una de las familias, realizada por personas pertenecientes a su misma etnia, en la medida que ello fuera posible.

Para agilizar la cuantificación y el tratamiento posterior de los datos, y para facilitar la tarea a los entrevistadores, se elaboró un protocolo de entrevista estructurado siguiendo los apartados que figuran en los objetivos generales.

2.1. Proceso de elaboración de la guía de la entrevista

Una vez delimitados los ámbitos del estudio, tal y como figuran descritos en los objetivos generales, se confeccionó un protocolo de entrevista, con una primera categorización de las respuestas.

Para realizar delimitación de los ámbitos que figuran en los objetivos y la categorización de las respuestas se consultaron estudios e investigaciones realizadas en el contexto familiar de las diferentes etnias, también se tuvo en cuenta el conocimiento directo de las pautas y costumbres culturales de las diferentes etnias por miembros del equipo.

2.2. Selección de los entrevistadores

Una vez confeccionado el protocolo se buscaron personas de las diferentes etnias que pudieran encargarse de realizar las entrevistas con las familias. Estas personas han actuado además como jueces para hacer una primera valoración del contenido de la entrevista y de la categorización realizada en el protocolo.

En el proceso de selección de estas personas se consideraron los criterios siguientes:

— La formación. Se han escogido personas con un nivel de formación universitaria que pudieran comprender la finalidad y objetivos de la investigación.

— El dominio de la lengua castellana además de la propia, con el objeto de que entendieran bien el contenido de las cuestiones que habían de plantear y supieran registrar la información recibida de una forma ágil en las casillas correspondientes.

— El grado de conocimiento de la cultura propia. Se han seleccionado personas que llevan aquí un tiempo, lo que les permite conocer la lengua, pero que han pasado la mayor parte de su vida en el país de origen. Son personas que conocen también la problemática que se puede derivar de la situación legal.

— Las capacidades personales en cuanto a iniciativa personal, facilidad para establecer contacto personal, fluidez de expresión, etc.

2.3. Valoración de la guía por los jueces

Se entregó el primer protocolo de entrevista al equipo que trabaja en el departamento de migración para que hicieran una valoración tanto del contenido como de la categorización. Simultáneamente se tuvo una entrevista colectiva con el grupo de personas seleccionadas de las diferentes etnias, para aplicarlo.

Tras efectuar una presentación de los entrevistadores se les explicó detalladamente los objetivos de la investigación y el contenido del protocolo. Se les entregó una copia del primer protocolo de entrevista y se fueron leyendo en voz alta cada una de las cuestiones detenidamente y cada uno de ellos iba comentando la oportunidad o no de su inclusión así como el tipo de categorizaciones realizadas.

Una vez revisado el protocolo en su totalidad se les pidió a cada uno que expusiera cómo realizaría la entrevista. Se hizo una pequeña escenificación por parejas y, a partir de ella, y de las apreciaciones realizadas anteriormente, el equipo de la investigación quedó encargado de revisar el orden de presentación de las cuestiones en el protocolo.

2.4. Formación de los entrevistadores

Tiene como principal objetivo que todas las personas de las diferentes etnias que participarán en la realización de las entrevistas comprendan bien los objetivos de la investigación y efectúen la entrevista siguiendo las mismas pautas. Para ello se efectuó una primera reunión con todas ellas, según se ha descrito en el punto anterior. La escenificación de la entrevista nos permitió comprobar en que medida habían captado el modo como se habría de desarrollar la entrevista.

El segundo paso fué entregarles a cada uno de los entrevistadores dos protocolos de entrevista para que realizaran una aplicación piloto y poder valorar con ellos las dificultades que habían encontrado, tanto en el proceso de la entrevista como en la categorización de las respuestas.

Se efectuó una segunda entrevista colectiva con todos ellos, en ella se fueron revisando las respuestas dadas por las personas entrevistadas de cada etnia y los problemas que habían surgido para categorizar las respuestas.

2.5. Aplicación piloto de la primera guía de entrevista

El objetivo de esta aplicación era doble, por una parte comprobar la adecuación de las categorizaciones de las preguntas, y por otra comprobar su utilidad para anotar las respuestas de las familias por parte de los entrevistadores, en este punto nos interesaba comprobar también el orden de presentación de las hojas de la guía, ya que las cuestiones estaban agrupadas por núcleos temáticos.

Se realizaron entrevistas con dos familias de cada una de las etnias seleccionadas en la muestra. La entrevista se efectuó en los propios domicilios de las familias, en algún caso estuvieron presentes los niños.

Los resultados de esta aplicación piloto nos permitieron puntualizar o matizar algunas de las preguntas categorizadas en el protocolo y añadir algunas cuestiones abiertas en los apartados que no figuraban originalmente. Se comprobó que no había dificultad por parte de las familias en responder a las cuestiones planteadas, aunque en algunos casos y debido a la situación legal, se mostró alguna reticencia al principio. Una vez que los entrevistadores explicaban el objetivo de la encuesta, y de la investigación las familias se mostraron colaboradoras.

2.6. Reestructuración de la guía de entrevista a padres

Con los datos aportados por la revisión efectuada por el equipo del Departamento de Migración, y los provenientes de la primera y segunda entrevista colectiva del equipo de investigación con los entrevistadores, el equipo revisó el contenido del protocolo, el orden de presentación de las cuestiones y la categorización de las respuestas. Se acordó que los entrevistadores recogerían las observaciones realizadas en cada aplicación en una hoja aparte al protocolo, además de anotar las observaciones específicas de algunas cuestiones que ya figuran en el protocolo. Se tendrá en cuenta en esta observación general:

- Actitud de la familia hacia las distintas cuestiones de la entrevista.
- Motivos de dejar alguna cuestión sin responder.
- Datos que quiera aportar la familia, sobre su proceso de integración, pautas culturales y problemática escolar de los hijos, etc, que no figuran explícitamente en el protocolo.

2.7. Presentación de la guía definitiva de la entrevista

La guía definitiva de la entrevista se presenta estructurada y ordenada según los siguientes apartados:

1. Descripción de las familias: datos que caracterizan el medio familiar
 - Aspectos /Variables
 - Datos descriptivos: País de origen. Lugar de nacimiento. Años de estancia en el país. Nº de países en los que han residido. Situación legal.
 - Inmigración: Motivos. Proyecto migratorio.
 - Datos familiares: Edades. Tipología de la pareja y la familia. Vivienda.
 - Nivel de formación: Estudios. Conocimiento de lenguas. Lengua habitual en casa.
 - Situación laboral: Sector laboral. Cualificación profesional. Situación laboral/no laboral. Tipo de Contratación. Formación Profesional/Ocupacional. Valoración del tipo de trabajo que desempeña.
2. Nivel de integración en el medio social próximo
 - Aspectos/Variables
 - Valoración de la cultura de origen: Nivel de integración en Asociaciones. Relaciones de amistad. Mantenimiento de costumbres propias. Relaciones informales.
 - Valoración de la cultura de acogida: Nivel de integración en asociaciones. Adquisición de otras costumbres. Relaciones informales. Participación en el centro escolar.
3. Valores y estereotipos culturales de los padres
 - Aspectos/Variables
 - Roles de autoridad y toma de decisiones: pautas educativas. Resolución de problemas de los hijos. Elección de matrimonio y futuro académico.

- Percepción de estereotipos en la cultura propia y de acogida: rasgos personales
- Diferenciación sexual de papeles: características personales, disciplina, actividades.
- 4. Datos de los hijos
 - Aspectos/Variables
 - Datos de los hijos: nombre, sexo, edad, curso escolar, lugar de nacimiento, residencia actual, años de escolarización en el país y fuera de él, situación legal, centro al que asisten.
- 5. Expectativas educativas de los padres
 - Aspectos/Variables
 - Funciones atribuidas a la escuela
 - Percepción sobre el nivel académico, la cualificación laboral y el tipo de trabajo de los hijos
- 6. Nivel de integración de los padres y actitudes hacia la escuela
 - Aspectos/Variables
 - Conocimiento: niveles de escolaridad, asociación padres, profesores de los hijos
 - Participación: asistencia a reuniones, celebraciones, fiestas.
 - Actitudes: lengua, costumbres, segregación, mantenimiento de las propias tradiciones
- 7. Problemas educativos detectados por los padres
 - Aspectos/Variables
 - Percepción de los padres sobre: actitud de los hijos hacia la escuela, dificultades académicas, nivel de los hijos con respecto a los compañeros
 - Relaciones de los padres con el centro: satisfacción, problemas y aspectos positivos.

2.8. Seguimiento de los entrevistadores

Las entrevistas con las familias se efectúan durante el mes de mayo. Semanalmente el equipo de investigación se encuentra con los entrevistadores para hacer un seguimiento del proceso. Se valoran en este encuentro las dificultades que tienen, los problemas surgidos en las familias y la forma cómo los han resuelto. Se les dan unas pautas comunes de actuación en las que se tienen en cuenta los objetivos de la investigación y la especificidad de reacción de cada una de las etnias en particular.

2.9. Análisis de datos de las entrevistas

Está previsto iniciarlo en el mes de septiembre. Para la valoración de los datos se preveen algunas sesiones conjuntas del equipo con los entrevistadores que se han ofrecido para colaborar en la interpretación de los resultados.

ESCALA DE ACTITUDES DEL PROFESORADO ANTE LA EDUCACIÓN MULTICULTURAL¹

F. Cabrera, J. V. Espín, M^a Ángeles Marín, M. Rodríguez Lajo y M. Torrado
Universidad de Barcelona

1. FINALIDAD

Se pretende conocer las actitudes del profesorado de Educación Primaria ante la educación multicultural. Su conocimiento permitirá constatar y valorar tanto situaciones individuales como de grupo y es importante para comprender cómo y por qué se actúa de una determinada manera en dichas situaciones. La medición de actitudes junto con los intereses y valores forman parte del dominio afectivo de las personas. Las actitudes son «predisposiciones aprendidas (no innatas) y estables (aunque pueden cambiarse) a reaccionar de una manera valorativa, favorable o desfavorable, ante un objeto». Desde el punto de vista de la medición, la actitud se considera una variable intermedia entre el estímulo u objeto de actitud y la respuesta o manifestación externa, de alguna manera mensurable.

2. METODOLOGÍA

Entre las distintas formas de medir actitudes se optó por realizarlo a través de la construcción de una **escala sumativa tipo Likert**, por las ventajas que presentaba respecto a otras formas o tipos de escalas diferentes.

La medición de actitudes mediante escalas se basa en la medición de las manifestaciones o respuestas de los sujetos ante opiniones referidas a creencias, sentimientos o conductas relativas al objeto medido. Las respuestas de los sujetos son reacciones valorativas que como tienen connotaciones afectivas y valorativas permiten inferir las actitudes. En el tipo de escala elegido, la suma de las respuestas a ítems supuestamente homogéneos sitúa al sujeto en la variable medida. La única suposición básica es que la respuesta evocada por cada ítem está en función de la posición del sujeto en el continuo de la variable medida: a más acuerdo (o desacuerdo según la dirección del ítem) posee más de la actitud medida.

1 El trabajo que se presenta en estas páginas debe enmarcarse dentro de la investigación más amplia descrita en otra comunicación: «*Diagnóstico de las diferencias étnicas y de los procesos desarrollados en la educación primaria. Percepción del profesorado*», por ello la bibliografía se recoge en dicha comunicación.

2.1. Proceso de elaboración de la escala

Las escalas de medición de actitudes tipo Likert están constituidas por juicios relativos al objeto de actitud a medir que expresen y constituyan una muestra representativa del conjunto de opiniones posibles que pueden manifestar las personas de interés para el estudio, según estén a favor o en contra de dicho objeto. En nuestro caso opiniones de interés que puedan manifestar los profesores respecto a la educación multicultural, en el ámbito de la Educación Primaria. Se siguieron los siguientes pasos:

2.1.1. Caracterización de la actitud y plan previo

Para la cobertura del espectro de opiniones que constituyen la actitud hay que conocer y seleccionar las dimensiones o aspectos que subyacen en ellas, pues orientan y dirigen la actitud. Hay que describir las características del sujeto alto o bajo en la actitud a medir en relación con dichas dimensiones. Estas descripciones equivalen a definir la actitud y sirven después de guía para la redacción de los ítems.

En nuestro caso el equipo investigador realizó una extensiva revisión bibliográfica, tanto de trabajos teóricos como empíricos, extrayendo cada investigador opiniones tanto a favor como en contra y comprobando en qué aspectos se repetían éstas con mayor frecuencia. Tras ello, se hizo una puesta en común y se seleccionaron las dimensiones que se vieron más relevantes, con sus matices correspondientes. **Las dimensiones seleccionadas** que debían cubrir los juicios fueron las siguientes:

- Efectos que se producen en los niños (en el grupo mayoritario, en los minoritarios y en el total) relativos a: socialización, aprendizaje, personalidad, interés o motivación, valores, disciplina.
- Efectos que se producen en el profesor, relativos a: su trabajo en el aula, relación con los alumnos, su formación, su personalidad.
- Efectos en el trabajo del aula, relativos a: programación de materias, materiales utilizados, contenidos que se trabajan, experiencias de aprendizaje.
- Papel de la escuela. Organización y tipo de enseñanza. Asignación de niños al aula y a la escuela.

Estas especificaciones organizadas por dimensiones constituyeron la guía para la elaboración de ítems.

2.1.2. Redacción de ítems: formulación de juicios

Los ítems se redactaron en forma de opiniones relativas a la educación multicultural con las que se puede o no estar de acuerdo.

Cada investigador formuló individualmente juicios que respondieran a las dimensiones elegidas, estando atentos en el momento de su elaboración para que su redacción no suscitara sesgos en las respuestas de los sujetos.

Días después se hizo una puesta en común examinando su relevancia, claridad, y que reflejaran opiniones y no hechos. Se seleccionó aquellos que por consenso parecieron mejores, teniendo en cuenta que tanto los que respondían a opiniones de personas a favor o en contra de la educación multicultural se distribuyeran atendiendo a las distintas dimensiones y que hubiera finalmente igual número de juicios positivos como negativos. La escala quedó constituida por 24 juicios.

Finalmente se atendió a otras cuestiones de la escala y su formato:

— Decisión sobre el número de respuestas posibles por ítem. Atendiendo a las consideraciones que deben tenerse en cuenta en la fijación del número de respuestas (edad de los encuestados, número de ítems y tipo de constructo) se decidió incluir cinco alternativas de respuesta.

- Organización de los ítems o juicios en la escala, para evitar sesgos en las respuestas debido al efecto de halo.
- Presentación del cuestionario de actitudes.
- Otras variables a medir: datos de identificación (sexo, años de profesión, años trabajando con niños de diferentes étnias).
- Instrucciones de respuesta.
- Recogida de opinión y valoración del cuestionario por los encuestados.

2.2. Aplicación piloto

Se realizó un estudio piloto para comprobar la adecuación de la escala y comprobar la eficacia de los ítems.

2.2.1. Muestra

La escala se aplicó a un grupo de 32 sujetos, constituido por alumnos de 5º curso de Pedagogía de la especialidad de Orientación, de horario de clase nocturno. Muchos de ellos tenían años de experiencia docente ejerciendo como maestros y algunos trabajando varios años con niños de diferentes etnias. Este hecho hizo que se considerara un grupo característico o parecido al del profesorado objeto de estudio, tanto para permitirnos analizar la adecuación de la propia escala como para poder hacer aportaciones sobre la misma, enjuiciándola. La aplicación se realizó en el mes de marzo.

2.2.2. Análisis de los resultados

Sobre el cuestionario de actitudes se analizaron dos tipos de datos:

A) Análisis cualitativo

El análisis cualitativo de los datos se realizó sobre las respuestas dadas por los sujetos a una pregunta abierta, formulada al final del cuestionario. En ella, se les pedía su opinión sobre el mismo en cuanto a presentación, contenido, estructura, aspectos no considerados o cualquier otra cosa que sobre el mismo desearan manifestar. Para ello se utilizó la técnica del «análisis de contenido». Se realizó el vaciado de las observaciones y se anotó las frecuencias de dichas observaciones. Como resumen podemos decir que aunque globalmente se consideró adecuado, se hicieron las siguientes sugerencias, que a su vez darán lugar a modificaciones:

— La alternativa de respuesta «Indiferente» no quedaba clara en cuanto a lo que significaba. Se modificará la escala sustituyendo esta alternativa por «indeciso», a la vez que se explicará el significado de cada una de las alternativas.

— Algunas preguntas resultaban repetitivas o reiterativas. Así el ítem 13 y 24 ó el 12 y 20 eran muy semejantes entre sí, se diferenciaban por pequeños matices. Aunque estos ítems se refieren al mismo tipo de efectos sobre el profesor y sobre la organización de los niños en la escuela o en la clase, se anulará uno de ellos y se clarificará la formulación en los otros.

— Algún ítem no resultaba todo lo claro que era de desear. Así, en el ítem 12 no quedaba claro lo que implicaba «sus propias escuelas». Se clarificará indicando «escuelas de su propia etnia».

— Algún sujeto indicó que sus respuestas podrían depender de la edad de los niños, por lo que se vio necesario dejar claro en la introducción que sus opiniones debían contextualizarse en la «Educación Primaria».

— Algún sujeto dudaba en los ítems relativos al profesor si debía responder según su caso particular o según lo que piensan en general los profesores. Se clarificará en la introducción que se trata de opiniones particulares.

B) Análisis Cuantitativo

Los datos cuantitativos se procesaron informáticamente utilizando el programa SPSS/PC+ versión 4.0., que aporta la siguiente información:

- Definición de la matriz de datos y presentación de la misma.
- Distribución de los sujetos de la muestra en las alternativas de los datos de identificación.
- Estadísticos descriptivos de cada ítem de la escala.
- Matriz de correlaciones entre ítems de la escala.
- Índice de homogeneidad de cada ítem, mediante la correlación ítem-total.
- Coeficiente de Fiabilidad de la escala como consistencia interna —mediante el coeficiente alfa de Cronbach—.
- Correlación ítem-total, eliminando la influencia del propio ítem y coeficiente de fiabilidad si se elimina cada uno de los ítems.
- Índice de discriminación de cada ítem.

Dado que la extensión de esta comunicación no nos permite dar todos los resultados, pasamos a continuación a presentar aquellos que son más interesantes para el objetivo de este estudio piloto. En concreto, la información referida al análisis de los ítems y fiabilidad de la escala.

Análisis de los ítems

En la tabla que aparece a continuación se presenta el índice de homogeneidad de cada ítem —mediante la correlación ítem-total— y el índice de discriminación. Los índices de discriminación se han obtenido contrastando las medias del grupo superior e inferior, mediante la prueba «t» de Student. En dichos grupos (g_1 = formado por el 50% inferior y g_2 = por el 50% superior) se trabajó con 26 sujetos, pues se eliminaron para este análisis aquellos sujetos que habían omitido su respuesta a algún ítem.

En la tabla referida al análisis de ítems, atendiendo a los niveles de significación, puede observarse, por ejemplo, como los ítems 6 y 11 guardan muy poca homogeneidad con el total y el contraste de medias tampoco es significativo. Además son los dos únicos ítems que si se eliminan se produce un aumento en el coeficiente de fiabilidad de la prueba y que guardan en mayor número de correlaciones negativas con otros ítems. Deben por tanto cambiarse. Otros ítems: 2, 3, 8, 9, 15, 18, 19, 20, 21, 22, 23 ó 24 resultan muy efectivos.

El grupo investigador fué revisando cada uno de los ítems, teniendo en cuenta tanto la información cuantitativa como la cualitativa.

Fiabilidad

Se obtiene un elevado coeficiente de fiabilidad como consistencia interna de la prueba. El coeficiente alfa de Cronbach es de 0,8279. Al determinar si el hecho de eliminar alguno de los ítems se produce cambios en dicho coeficiente se comprueba que todos ellos contribuyen a la consistencia interna, excepto el 6 y el 11 que de eliminarse aumentaría el coeficiente a 0,8384 y 0,8398 respectivamente.

ANÁLISIS DE ÍTEMS

ítem	I. de Homogeneidad	I. Discriminación	
1	0,4852**	2,32*	Si
2	0,5939**	2,41*	Si
3	0,5314**	2,51* Si	
4	0,4643**	1,62	No
5	0,5196**	1,91	No
6	0,1581	0,38	No
7	0,5004**	1,41	No
8	0,6225**	2,63*	Si
9	0,5618**	2,01	No
10	0,5214**	2,01	No
11	0,1244	0,19	No
12	0,3505*	1,39	No
13	0,3907*	2,71*	Si
14	0,5056**	1,87	No
15	0,5940**	2,38*	Si
16	0,4047*	1,76	No
17	0,3949*	1,21	No
18	0,7196**	3,83**	Si
19	0,3678*	2,47*	Si
20	0,5036**	2,08*	Si
21	0,5363**	3,69**	Si
22	0,4348**	2,65*	Si
23	0,4048*	2,60*	Si
24	0,4569**	2,83**	Si

* significación al 5%

** significación al 1%

2.3. Reelaboración de la escala

Tras el análisis cualitativo y cuantitativo de los resultados se realizaron las siguientes modificaciones de la escala:

Modificaciones en los ítems

A excepción de los ítems que resultaron ser efectivos, y que se mantuvieron igual, se realizaron distinto tipo de modificaciones. Se anuló el ítem 6 que se cambió por otro y se modificó el 11. Este ítem no se anuló porque la dimensión a que respondía se consideró relevante. En otros ítems se hicieron pequeñas modificaciones encaminadas a una mayor claridad. Así, por ejemplo, en el ítem 15 se cambió «aprenden más» por «mejor rendimiento académico»; el ítem 1 hacía referencia simultáneamente a dos tipos de efectos en los niños (apertura y solidaridad) y se modifica para que aluda a un único tipo de efecto; en el ítem 12, ya comentado en el análisis cualitativo, se cambia «sus propias

escuelas» por «escuelas de su propia étnia» etc... También en los ítems donde aparecía alusiones a las diferencias de «raza» «étnia» y «cultura» se mantuvo sólo la étnia.

Una vez modificados los ítems se comprobó de nuevo que estuvieran bien representadas las dimensiones elegidas del constructo, así como su organización y disposición en la escala. El número total de ítems fue de 24, siendo igual el número de positivos que de negativos.

Cambios en el formato

— Se cambia la forma de presentar las alternativas de respuesta a los ítems. En lugar de situar las casillas de respuesta con sus descripciones respectivas debajo de cada ítem, se sitúan a la derecha de los mismos, apareciendo la descripción de su significado en la parte superior de la hoja. La presentación gana claridad y no es monótona.

— Se añade una introducción a la escala, en donde se expresa la necesidad de conocer su opinión como profesores sobre la educación multicultural, contextualizada en el ámbito de la Educación Primaria. También se les indica que no se preocupen si las situaciones no matizan su contenido pues interesa su opinión global.

— En cuanto a los datos de identificación, se cambia una de las alternativas de los años de profesión (pues no tiene sentido incluir el valor cero) y se añaden nuevos datos: tipo de centro y ubicación del mismo según sea rural, urbano o suburbano. Los datos de identificación se presentan enmarcados para diferenciarlos bien de la escala, y aparecen en la primera página, tras la introducción.

— El protocolo de la escala queda definitivamente presentado en forma de cuadernillo. Las instrucciones de respuesta se presentan por escrito y la escala se redacta en lengua catalana.

La aplicación de la escala definitiva se inicia en el mes de mayo y está en fase de realización.

EL MAPA COGNITIVO DE LAS PROFESIONES

M^ª Teresa Muñoz Sastre
U.N.E.D.

Según Gottfredson (1981), los adolescentes construyen progresivamente un espacio de profesiones «aceptables» en base a cuatro criterios generales: género, criterio biológico; estatus, criterio social; campo de interés, criterio psicológico; y, accesibilidad a la formación y a los empleos, criterio basado en el principio de la realidad. Evolutivamente hablando, género es aparentemente el primer criterio en ser tenido en cuenta. La significación social de sex-tipo es claramente percibida desde los 7-8 años. Por esta razón, el espacio de las profesiones a esta edad puede vivirse como dividido en dos partes. Por un lado, las profesiones compatibles con el criterio de sex-rol, por otro las profesiones inaceptables (típicamente masculinas para las chicas y típicamente femeninas para los chicos). Estatus social es el segundo criterio que se tiene en cuenta y comienza a ser percibido desde los 10 años. Chicos y chicas progresivamente tienden a restringir la zona de alternativas aceptables a las profesiones compatibles con estatus social. Dada la relativa independencia de los criterios género y estatus social, el espacio de las profesiones puede aparecer como bidimensional aproximadamente a los diez años de edad. El criterio psicológico, dominio de interés, es el tercer criterio que se desarrolla. Los intereses y aspiraciones propios son tenidos en cuenta hacia los 14 años. Progresivamente, una nueva restricción de la zona de alternativas aceptables aparece en función de la percepción de sí mismo con características únicas. El criterio de accesibilidad hace su aparición en un momento en que pueden haberse ya tomado decisiones más o menos irreversibles. Es un criterio del tipo costo-valores. Ciertas profesiones que han sobrevivido a las tres restricciones previas parecen de difícil acceso por tanto tienden a desaparecer de la zona de las alternativas aceptables.

Según Gottfredson, esas restricciones o extensiones progresivas tendrían lugar en un espacio mental de dos dimensiones que ella llama mapa cognitivo de las profesiones. Género y Estatus social serían las dos dimensiones que definen este espacio. Los campos de interés (Holland, 1985), que cristalizan todo lo relativo a los intereses profesionales clásicos, a los valores, a las convicciones referentess a ciertas capacidades o habilidades específicas, podrían estar localizadas en el interior mismo de este espacio. El campo Investigativo, por ejemplo, en que las coordenadas que representa (Cirujano, Psiquiatra, Físico, Juez Federal, etc.) son las más elevadas en términos de estatus social y se sitúan al nivel de los dos tercios del eje Masculinidad-Feminidad. El campo Realista es en el que las coordenadas de las profesiones que representa (Minero, Pescador, Mecánico de automóviles, Cocinero,...) son las más bajas en términos de Estatus social, y se sitúan más cerca del polo Masculinidad del eje Género. El campo Convencional, último ejemplo, es en el que las coordenadas de las profesiones que representa (Azafata, Secretario, Empleado de banco, etc.) están igualmente entre las más bajas en términos de Estatus social pero se sitúan esta vez muy cerca del polo Feminidad del eje Género.

Las cuestiones que nos hemos planteado se refieren a este mapa cognitivo de las profesiones y más concretamente a su dimensionalidad en el caso de los adolescentes de 14 años, es decir, a una de las edades que Gottfredson considera como edad clave en cuanto a la construcción del mapa mental de las profesiones puesto que corresponde con el comienzo de la integración dentro del espacio de las profesiones de las consideraciones relativas al sí mismo, a los gustos, los intereses, los valores. ¿En general, son suficientes dos dimensiones para estructurar el conjunto de las representaciones profesionales de los adolescentes de esta edad? Si es así, ¿éstas son Género y Estatus social? ¿Qué relaciones mantienen éstas con los valores?

MÉTODO

Sujetos. La muestra se compone de 735 adolescentes españoles realizando 8º de E.G.B. en 1988-89. Tienen una media de edad de 14 años. Todos viven en la provincia de Guadalajara.

Material. Los cuestionarios contienen un conjunto de 21 profesiones elegidas en función de tres criterios: nivel de cualificación, dominio de interés y grado de atracción. Fueron definidos tres tipos de nivel de cualificación: Formación Profesional, B.U.P. y Estudios Superiores Universitarios. Se tuvieron en cuenta los campos de interés clásicos: Literario, Artístico, Científico, Deportivo, Altruista, Naturaleza, Técnico, Administrativo y Empresarial. Además, solamente fueron retenidas las profesiones que en un pretest no fueron ni sistemáticamente elegidas ni sistemáticamente rechazadas por los estudiantes de 8º de E.G.B. Las 21 profesiones fueron apareadas de la siguiente manera: Dentro de una urna se introdujeron ochenta y cuatro papeletas con el nombre de una de las 21 profesiones, habiéndose escrito cada una de ellas cuatro veces. A continuación las papeletas se fueron sacando al azar. Los primeras dos profesiones sacadas formaron el primer par, las dos segundas el segundo par y así sucesivamente. Esto nos dio un total de 42 pares de profesiones, apareciendo cada profesión exactamente cuatro veces. El material también incluyó un conjunto de 12 consignas referentes a 12 características distintas utilizadas para juzgar esas profesiones. El material fue previamente testado sobre una población de adolescentes españoles (Muñoz Sastre y Mullet, 1990, 1991; ver también 1992).

Procedimiento. Primero se presenta a los sujetos un primer cuestionario con los 42 pares de profesiones y se les pide que designen, en cada par, la profesión que parece ofrecer los ingresos económicos más elevados (o las mayores posibilidades de contactos sociales, o las mayores posibilidades de tiempo libre, etc...). A continuación se les presenta un cuestionario idéntico, consistiendo esta vez la consigna en designar, de las dos profesiones, la que parece ofrecer las mayores posibilidades de contactos sociales, etc. Cada sujeto ha rellenado cuatro cuestionarios sacados al azar de los doce posibles. El tiempo es libre. La duración total de pasación es de unos 75 minutos. Antes de efectuar sus designaciones, los alumnos tenían la posibilidad de informarse sobre las profesiones que no conocían preguntando a la experimentadora (yo misma). Los cuestionarios son en general completos (pérdida inferior a 1%).

RESULTADOS

De cada sujeto hemos obtenido datos procedentes de cuatro series de 42 designaciones. Por cada sujeto, fueron comparadas las respuestas a las series dos a dos y puntuadas 1 cada vez que dos designaciones conciden. Cada vez que las respuestas de cada comparación no coincidía se puntuaba 0. Se calcularon coeficientes Phi (entre cada par de cuestionarios. La fórmula fue $\Phi = [(Puntuación\ total - 21)2 / 21^2]^{1/2}$. Esto dió seis coeficientes Phi por cada sujeto. Las medias de esos coeficientes han sido objeto de diferentes tratamientos. Cada uno de esos valores medios ha sido calculado sobre un subconjunto de sujetos cuyo efectivo es aproximadamente 67 (739 sujetos x 6 / 12 x 11/2).

TABLA 1

FACTORES	I	II	III	IV
1-Feminidad				.83
2-Prestigio	.78			
3-Salario	.75			
4-Promoción	.48			
5-Exterior		.55		
6-Tiempo lib.		.61		
7-Manual		.74		
8-Iniciativa		.66		
9-Encuentros		.56	-.34	.46
10-Acc. costo	.63			
11-Acc. éxito	.40			
12-Salidas			.63	

Ha sido realizado un análisis factorial en componentes principales sobre los 66 valores de los coeficientes Phi. Se han revelado cuatro factores presentando un valor propio superior a 1. Se han conservado esos cuatro factores y sobre ellos se ha aplicado una rotación varimax. Los coeficientes de saturación se presentan en la tabla 1. El Factor I explica alrededor del 22% de la varianza total explicada. Este factor satura esencialmente los determinantes de tipo Estatus Social así como los dos determinantes de accesibilidad. Podemos interpretarlo como el factor Estatus social. Si las variables de accesibilidad hubieran sido introducidas con un signo negativo, ese primer factor habría opuesto Estatus social y Accesibilidad. Hay que recordar que la consigna para estos dos determinantes es la inversa de la que se ha utilizado para los otros determinantes (designar la menos accesible). El factor II explica 15% de la varianza total explicada. Satura cuatro de los cinco determinantes de tipo Valores. Se puede, por tanto, interpretar como un factor que reagrupa lo esencial de los valores que hemos introducido. El factor III explica 12% de la varianza total explicada. Satura esencialmente Manual y Salidas (posibilidades de trabajo). Podría interpretarse como un factor Accesibilidad a los empleos, completamente independiente del factor Estatus social y de los dos determinantes de accesibilidad. La asociación Manual y Salidas deja entender que este factor es esencialmente «profesional». Podría corresponder con la división del trabajo entre cuellos azules, numerosos pero no necesariamente los peor remunerados, y cuellos blancos, intelectuales, menos numerosos y no necesariamente los mejor remunerados. Por último, el factor IV explica igualmente 12% de la varianza total explicada. Satura sobre todo Feminidad. Se interpreta claramente como el factor independiente Género propuesto por Gottfredson. El determinante Encuentros, asociado a este factor, viene a reforzar la interpretación.

Discusión

Los resultados del análisis factorial tenderían, en parte a confortar las proposiciones de Gottfredson y en parte a negarlas o al menos a complementarlas. Existiría un factor Género independiente del factor Estatus social, constitutivos del mapa mental de las profesiones, pero ciertos valores tienden

también a formar un factor independiente, nuestro factor II. Uno al menos de esos valores se encuentra ligado al factor Género (Encuentros). Finalmente, la emergencia de un cuarto factor, independiente de Estatus social, y evocando fuertemente la división de trabajo Manual-Intelectual, puede ponerse en relación con resultados más antiguos, los de Reeb (1974). Por tanto, hay que considerar que la reducción a dos dimensiones del mapa cognitivo de las profesiones constituye una proposición que se puede poner en causa. El concepto de mapa cognitivo de las profesiones en el caso de los adolescentes, jóvenes o menos jóvenes, convendría sustituirlo por el concepto de espacio cognitivo, término que evoca una dimensionalidad, una riqueza, de un orden más elevado. Esta concepción más rica tiene como consecuencia el hacer más coherente con las leyes ordinarias de la geometría, ciertas proposiciones complementarias de Gottfredson sobre lo que ella llama el proceso de compromiso. Ante la necesidad de tener en cuenta la realidad, es decir, en los diferentes momentos de elección efectivos más o menos definitivos, Gottfredson avanza la hipótesis siguiente: «people will tend to sacrifice interests in field of work to maintain sextype and prestige, and to some extent will sacrifice prestige level for sex-type if that is also necessary» (p. 572). Pero cómo modificar, en un espacio de sólo dos dimensiones, sus intereses o valores sin modificar al mismo tiempo el nivel de estatus social y/o el nivel de Género. Eso no llega a ser «geométricamente» posible que en el caso de que se considere un sistema de representación con más de dos factores, de los cuales uno puede interpretarse como correspondiendo a los Valores profesionales y a propósito de los cuales pueden intervenir variaciones independientes de Género y Estatus social.

REFERENCIAS

- GOTTFREDSON, L. S. (1981): Circonscription and compromise: A developmental theory of occupational aspirations (Monograph). *Journal of Counseling Psychology*, 28, 545-579.
- HOLLAND, J. L. (1985): *Making vocational choices: A theory of vocational personalities and work environments*. Englewood Cliffs, NJ: Prentice Hall.
- MUÑOZ SASTRE, M. T. y MULLET, E. (1990): Los determinantes de las preferencias profesionales de los adolescentes. *Psychological Assessment*, 6, 155-170.
- MUÑOZ SASTRE, M. T. y MULLET, E. (1991): Un procedimiento educativo para permitir a los adolescentes una toma de conciencia de algunos de los determinantes de las preferencias profesionales. *Bordón*, 43, 67-78.
- MUÑOZ SASTRE, M. T. and MULLET, E. (1992): Occupational preferences of Spanish adolescent in relation to Gottfredson's theory. *Journal of Vocational Behavior*, 40, 306-317.
- REEB, M. (1974): The perception of occupational structure: An intervening variable in vocational behavior. *Journal of Vocational Behavior*, 4, 125-137.

ESTIMACIÓN DEL VOCABULARIO TOTAL CONOCIDO POR UN ESCOLAR

A. Suárez

Universidad de Santiago

P. Meara

University College, Swansea

I. INTRODUCCIÓN

Entre los múltiples estudios cuantitativos referidos al vocabulario, existen aquellos que nos permiten estimar el número total de palabras «conocidas» por una determinada persona. Nos permiten afirmar, por ejemplo, de dos escolares de 10 años de edad, que uno conoce unas 5.000 palabras y el otro, unas 15.000, del total de palabras de un universo léxico prefijado.

El número total de palabras conocidas por una persona es una información mucho más rica que la obtenida mediante los tests habituales de vocabulario —con frecuencia, de elección múltiple—, que sólo nos dicen el lugar que ocupa el sujeto dentro del grupo normativo. La comparación de las diferencias del tamaño total del vocabulario en momentos sucesivos de la vida de una persona nos permite establecer el ritmo de crecimiento del vocabulario, dato del mayor interés, sobre todo para no caer en el error corriente de la subestimación del tamaño total del mismo (Seashore, 1949; Crystal, 1990) y en las indeseables consecuencias que de esta subestimación se derivan (Por ejemplo, la de sobrevalorar la importancia de la enseñanza directa de un vocabulario «tipo», como se hace en Casanova y Rivera, 1989, en lugar de insistir en los mecanismos indirectos de adquisición, que podríamos resumir con el constructo «volumen de experiencias lingüísticas» (Nagy & Anderson, 1984).

Ha habido diversos intentos de contestar a la pregunta de cuántas palabras conoce una determinada persona. Por un lado, tenemos estudios de tipo biográfico (Leopold, 1949; Nelson, 1973; Wagner, 1985, etc.), que registran las palabras que utiliza activamente el sujeto, actualmente mediante el uso de técnicas radio-microfónicas. Por otro lado, tenemos los estudios basados en cálculos inferenciales: se obtiene una muestra de los vocablos de un diccionario por el sistema de catas, se averigua cuántos de esos vocablos conoce el individuo y, finalmente, se hace la estimación inferencial de cuántas palabras conoce de las del diccionario de partida (Smith, 1926; Seashore & Eckerson, 1940; García Hoz, 1946; Diack, 1975, entre otros). Ambos tipos de estudios tienen limitaciones obvias.

II. NUESTRO TRABAJO

1. **Naturaleza de los tests.** Nuestra investigación se inscribe entre los trabajos que tratan de estimar el número total de vocablos conocidos por una persona a través de cálculos inferenciales. La aportación más novedosa de nuestro intento quizá consista en la naturaleza de los tests utilizados.

Cada test está formado por 40 palabras y 20 pseudopalabras. Como ilustración, presentamos los seis primeros estímulos de uno de los tests:

- | | | |
|-------------------|-------------------|-------------------------|
| 1 () narigudo | 2 () remaquir | 3 () encarecimiento |
| 4 () agenio | 5 () roscalde | 6 () regente |

Las 40 palabras pertenecen a una muestra aleatoria constituida por el 10% del total de unas 22.000 entradas que componen el Diccionario Escolar Etimológico (García Hoz, 1989, 17ª edición, actualizada y corregida). En esas 40 palabras están representadas proporcionalmente las palabras de cuatro estratos diferentes, correspondientes al «grado de dificultad de aprendizaje que presenta cada palabra», según el juicio emitido por tres profesores de E.G.B. Las pseudopalabras tienen características fonotácticas y grafotácticas similares a las palabras de la muestra y han sido generadas a partir de estas.

Al sujeto se le pide únicamente que escriba S (=sí) dentro del paréntesis si sabe lo que significa la palabra o que escriba N (=no) si no sabe lo que significa o no está seguro. Es una tarea motivadora para los sujetos, que realizan en unos tres minutos. Esta última característica permite aplicar varios de estos tests en el mismo tiempo que se administra un test convencional de vocabulario, lo que hace posible realizar las inferencias proporcionales a partir de más datos.

Las 20 pseudopalabras constituyen una especie de escala de honestidad, que, teniendo en cuenta la teoría de detección de señales, nos permite hacer una corrección del número de palabras realmente conocidas cuando el sujeto rodea como conocidas algunas pseudopalabras.

No podemos entrar aquí a fondo en el tema de **qué significa conocer una palabra** (AILA Review dedica un monográfico en 1989 al tema). Sin duda, nuestros tests sirven para averiguar, cuando menos, si el sujeto sabe distinguir lo que es y no es una palabra castellana, y es difícil imaginar un aspecto del conocimiento de una palabra más fundamental que este.

2. Hipótesis y diseño. Tratamos de someter a comprobación la hipótesis de que administrando a un sujeto un corto número de estos tests —independientemente de cuáles sean— siempre obtenemos la misma estimación de su vocabulario total. Si estimamos el tamaño del vocabulario a partir de un solo test, probablemente obtendremos una variación considerable entre las estimaciones de un test a otro; pero si lo estimamos a partir de la media obtenida en varios tests, reduciremos esta variación. La cuestión que tratamos de resolver es averiguar cuántos tests tenemos que administrar a un sujeto para obtener una estimación estable del tamaño de su vocabulario.

El diseño es de un factor con medidas repetidas y la técnica utilizada en el análisis de los datos, ANOVA. Podremos seguir manteniendo nuestra hipótesis en la medida que no hallemos diferencias significativas en las estimaciones medias obtenidas a partir de un cierto número de tests. Este resultado tiene que ir acompañado de altas correlaciones entre las medias de los grupos de tests en cuestión, para asegurarnos de que las estimaciones tampoco varían a nivel individual.

3. Administración de las pruebas. Se realizaron dos administraciones, ambas a alumnos de 6º de E.G.B., la primera en un Colegio Público de Alcobendas (Madrid) y la segunda en un Colegio subvencionado de Santiago de Compostela. En Alcobendas se administraron 5 de nuestros tests durante una sesión de la clase de Lengua y otros 5 una semana después. A todos los alumnos se le administraron los mismos tests y en el mismo orden. En Santiago, se administraron 4 tests en cada sesión y el orden de los tests para cada sujeto fue determinado al azar. En ambos casos entre las dos administraciones se aplicó a los escolares el subtest de vocabulario del test T.E.A.-1 (versión 1990).

4. Resultados. En la administración a los escolares de Alcobendas, la única diferencia no significativa obtenida —de las varias combinaciones comprobadas— se obtuvo cuando se compararon las estimaciones de los cinco tests administrados el primer día con las de los otros cinco tests aplicados una semana después: $F(1,105) = 1.61$. La correlación entre estos dos grupos de tests —«antes» y

«después»— fue $r = .7794$. Todas las correlaciones con el criterio (el subtest del T.E.A.-1 citado) resultaron positivas y significativas, incluso las de los tests considerados separadamente, que oscilaron entre $r = .25$ y $r = .51$, $p < .01$.

Al analizar los datos de la administración de Santiago, tampoco se obtuvieron resultados que siguieran una pauta clara: en algunos casos, la media en tres o cuatro tests resultó suficiente para obtener estimaciones del vocabulario total estables; en otras, no. Las correlaciones entre las medias de dos tests distintos van desde $r = .45$ a $r = .74$, $p < .001$. Las correlaciones entre cada uno de los tests y el criterio están entre $r = .20$ y $r = .43$, $p < .05$. La estimación de la media del vocabulario total para el grupo de alumnos de Santiago resultó ser $X = 10.306$, con una $s = 1.971$. Por supuesto, este vocabulario total ha de considerarse como un mínimo, ya que se trata de una estimación hecha a partir de un diccionario de sólo unas 22.000 entradas (Recordemos que el diccionario de la R.A.E. tiene cerca de 100.000 entradas).

III. CONCLUSIONES Y PROSPECTIVA

Los resultados obtenidos no constituyen una base empírica suficiente para poder sostener que nuestros tests sean instrumentos apropiados para estimar el vocabulario total conocido por un escolar. Sin embargo, existen varias razones que nos llevan a pensar que no debemos abandonar el intento: algunos de los resultados son esperanzadores, una simple inspección de los protocolos revela pautas claras en la mayoría de los sujetos, el intento en su conjunto parece obedecer a un planteamiento lógico, el propósito que perseguimos es decididamente relevante...

Nuestro programa inmediato de trabajo, guiado por la hipótesis explicitada anteriormente, tiene estos objetivos: mejorar las condiciones en que se realizan las pruebas y las instrucciones dadas a los sujetos (tratamos de disminuir la tasa de pseudopalabras rodeadas como conocidas); conseguir un modelo más satisfactorio para efectuar las correcciones de acuerdo al número de pseudopalabras rodeadas; administrar las pruebas a otros grupos de edad, etc.

BIBLIOGRAFÍA

- CASANOVA, M^a A., RIVERA, M. (1989): Vocabulario básico en la E.G.B., I-III. Madrid: M.E.C./ Espasa Calpe.
- CRYSTAL, D. (1989): Teaching Vocabulary: The Case for a Semantic Curriculum. En K. MOGFORD-BEVAN & J. SADLER: *Child Language Disability, V. II*. Clevedon, Avon: Multilingual Matters Ltd.
- DIACK, H. (1975): Standard Literacy Tests. St. Albans, Herts.: Hart— Davis Educational Ltd.
- GARCÍA HOZ, V. (1946): Evolución cuantitativa del vocabulario en escolares de nueve a dieciocho años. *Revista Española de Pedagogía*, 16, oct-dic., t. IV.
- JUSTICIA, J. F. (1984): Determinación y Análisis del Vocabulario del Niño de 6 a 10 años. Madrid: Memoria del Proyecto de Investigación 4.107/79, CAICYT.
- NAGY, W. E. & ANDERSON, R. C. (1984): How many words are there in printed school English? *Reading Research Quarterly*, v. XIX, 3.
- MEARA, P. & BUXTON, B. (1987): An alternative to multiple choice vocabulary tests. *Language testing*, 4, 2, 142-154.
- SEASHORE, R. H. & ECKERSON, L. D. (1940): The measurement of individual differences in general English vocabularies. *J. of E. Psychology*, 31, 14-38.

ANÁLISIS DEL TEST PMS DE RAVEN EN LOS ESCOLARES DE GALICIA

Eduardo Abalde Paz y J. Miguel Muñoz Cantero
Universidad de La Coruña

Todos sabemos que para medir la inteligencia es necesario tener en cuenta el instrumento que se va a utilizar. Por otro lado, el hecho de llevar a cabo un estudio de índole psicométrico, centrado en una población, en este caso la gallega, viene a justificar la necesidad cada vez mayor de posibilitar la comparación de las puntuaciones obtenidas por sujetos en distintos grados de escolarización, con un marco de referencia propio a las características implícitas a la propia Comunidad a la que pertenecen.

Nuestro trabajo se ha centrado en el estudio y análisis del test Matrices Progresivas, Nivel Superior, de J. C. Raven, en la población correspondiente a las cuatro provincias gallegas.

Un buen nivel de educación es un aspecto básico y fundamental para sostener la evolución y el progreso de una región, pero en Galicia no se puede hablar de una situación óptima. Las causas de estas deficiencias están relacionadas, en parte con el «tradicional atraso» de la región respecto al resto de España, y también, con su particular estructuración de espacio físico y humano. Los contrastes sociales y espaciales entre el mundo rural y urbano refuerzan las desigualdades; la infraestructura educativa; la repartición de los centros de EGB; el problema de la escolarización en el interior, hacen de esta Comunidad una unidad muy peculiar y susceptible de ser estudiada como tal.

Para realizar nuestro estudio, nos planteamos qué tipo de variables que debiéramos estudiar y que nos reflejaran características propias de la población objeto a estudio. En este sentido, se consideraron, aparte del sexo y la edad, el hábitat, en el que tuvimos en cuenta no sólo los asentamientos urbanos y rurales, sino también los pesqueros. Otras fueron el nivel educativo y profesional del padre, en tanto que siendo variables clásicas en estudios educativos, no por ello dejan de ser importantes para conocer si influyen o no en los resultados que los hijos obtienen en el test.

Utilizamos una muestra de 3.128 alumnos de EGB (sexto, séptimo y octavo); BUP y COU, y FP, pertenecientes a diferentes centros educativos de la Comunidad Autónoma de Galicia, utilizando el muestreo por conglomerados. Los criterios utilizados para la selección de centros responden a las siguientes variables estratificadoras:

- a.— Tipo de enseñanza: EGB, BUP, COU y FP.
- b.— Hábitat: Urbana, rural, pesquera.
- c.— Provincia: La Coruña, Lugo, Orense, Pontevedra.

A todos los sujetos que componen la muestra se les aplicó el test de matrices progresivas de Raven —Nivel Superior—. Primero se entregó la Serie I con el objeto de que el alumno conociese el método de trabajo necesario para la correcta utilización de la prueba dejándoles tiempo suficiente para que pudieran comprender el proceso en que se basa el PMS. A continuación se repartieron los cuadernillos de la Serie II.

Eliminados los protocolos incorrectos y codificados los resultados, se calcularon los diferentes estadísticos (medias, desviaciones típicas, número de sujetos para cada caso).

Los datos, una vez tratados estadísticamente, se analizaron según tres aspectos:

a.— análisis del test.

1.— Coeficiente de fiabilidad.

2.— Índice de dificultad Vs. facilidad.

3.— Índice de consistencia de cada ítem con el total.

b.— Estudio global del test, según las variables: Sexo, hábitat, Nivel educativo y Profesión del Padre.

Entre las diversas técnicas utilizadas en nuestra investigación han sido:

a.— Para la fiabilidad la correlación de Pearson y la de Kuder-Richarson (KR-20).

b.— Para el índice de dificultad, los porcentajes.

c.— Para índice de Consistencia Interna, la correlación biserial puntual y KR-20.

d.— Para la matriz de correlaciones inter-ítems, la correlación de Pearson.

e.— Para las diferencias entre grupos se utilizó primero n análisis de varianza simple y después la prueba t.

Los análisis psicométricos efectuados nos permiten señalar la alta fiabilidad del PMS de raven, por cuanto los índices obtenidos alcanzan valores suficientemente significativos. A nivel global de la muestra se alcanza una correlación de .8792 por el método de mitades y de .8683 con el de consistencia interna (KR20). Diferenciando según el sexo, hemos llegado a resultados muy similares.

Por edades, los valores de la fiabilidad son el general altos, van desde .8039 hasta .9320, excepto en 11 años que se alcanza únicamente una correlación de .7705, lo que confirma que ciertamente el test no es aconsejable para esta edad.

Tenemos que afirmar también que la fiabilidad por cursos es satisfactoria en todos los casos, puesto que los valores oscilan dentro de la banda de .80 y .90, considerada por diversos autores como muy aceptables. Es precisamente en sexto de EGB donde se da el índice más bajo .7940, que es justo el curso que corresponde a la edad de 11 años.

En lo que a validez se refiere, podemos indicar que, en general, los valores obtenidos son satisfactorios. No obstante, cabe destacar los hallados con el test D-48 (.68) y con el factor R del PMA (.37), y con las calificaciones escolares (.36), podemos afirmar que los valores alcanzados están dentro de los límites normales de validez.

En cuanto a los índices de dificultad Vs. facilidad de los ítems se puede afirmar que, en conformidad con las normas de construcción de pruebas psicológicas, se advierte una tendencia progresiva de dificultad en los diferentes elementos del test. En la muestra total, el 42% de los ítems oscilan entre I.D. (76-86/24-4) y el 58% está en los límites de (75-25). No existen diferencias entre los I.D. si se analiza por separado, en la muestra total, hombres y mujeres.

Del análisis que se ha realizado considerando las diferentes variables, diremos que el test resulta más difícil a los alumnos de zona rural, puesto que es donde menos ítems muy fáciles aparecen y mayor número de elementos muy difíciles.

Según la profesión del padre, podemos afirmar que la facilidad del test está en relación directa con la importancia de la categoría profesional, siendo más difícil en la última categoría.

Por cursos, tenemos que destacar, en primer lugar, a sexto de EGB y primero de FPI, para los que el test no parece apropiado, puesto que no hay ningún elemento que pueda considerarse muy fácil para ellos, y en cambio destaca un 47% de ítems muy difíciles. En los otros cursos podemos afirmar que los ID siguen, en general, las normas de construcción de pruebas psicológicas, resaltando en BUP, COU y FPPI, excepto primero, un 30%, por término medio, de ítems muy fáciles, así como un 25% de muy difíciles, y en los restantes cursos la distribución es de 8% y de 35 respectivamente.

Podemos concluir que la dificultad disminuye a medida que se asciende en curso y nivel educativo, teniendo presente que todo FPI y primero de FPII son equivalentes a la Educación General Básica, y que los demás cursos de FP de segundo grado se corresponden con el Bachillerato.

Del análisis global que hemos realizado con los resultados de la muestra total, en el que hemos considerado cada variable aisladamente podemos obtener varias conclusiones:

a.— No existen diferencias estadísticamente significativas en cuanto a hombres y mujeres en la Serie II del test; la probabilidad que obtenemos con la prueba t es $p=.495$; sin embargo, sí existen en la Serie I, con una $p=.002$, siendo la diferencia a favor de los hombres.

b.— En cuanto a la variable hábitat, tenemos que afirmar que existen diferencias según el análisis de varianza realizado entre los diversos tipos, y en las dos series. En la I la probabilidad es $p=.000$ entre los tres grupos, con valores más altos en los de zona pesquera y urbana. En la Serie II no existen diferencias entre estas dos últimas, sí entre éstas y la rural $p=.000$. Tenemos que afirmar que el hábitat urbano y pesquero pertenecen a la misma población en cuanto a la inteligencia media por el PMS. Los de zona rural son los que obtienen las puntuaciones más bajas en las dos series.

c.— Al estudiar la variable nivel educativo del padre, el análisis de varianza aplicado nos da diferencias significativas en ambas series con una $p=.0000$. Analizando los resultados, encontramos que los valores más altos los obtienen los hijos cuyos padres tienen Bachiller Superior; los más bajos, los de Estudios Primarios; y, por último, diremos que los de Bachiller Elemental alcanzan mejores puntuaciones que los Estudios Medios.

Tenemos que concluir que el mayor nivel educativo del padre no influye en los resultados que los hijos obtienen en el PMS.

d.— No sucede lo mismo con la variables profesión del padre, donde los estadísticos más altos se dan en categorías profesionales altas y van en orden descendente hasta categorías más bajas.

BIBLIOGRAFÍA

- ABALDE PAZ, E. y GARCÍA LÓPEZ (1987): *Análisis y Baremación del P.M. 56 Matrices progresivas en la población escolar de Galicia*. La Coruña. (eds.).
- ABALDE PAZ, E. y MUÑOZ CANTERO, J. M. (1992): «El PM-56 Matrices Progresivas —Escala General— en la Población escolar de la Comunidad Autónoma de Galicia». *Rev. Psicopedagogía* 6-7, 205-219.
- ABALDE PAZ, E. y MUÑOZ CANTERO, J. M. (1993): *El test PMS de Raven y los escolares de Galicia*. La Coruña. Servicio de Publicacions.
- BINGHAM, W. C.; BURKE, H. R. & MURRY, S. (1966): «Raven's progressive matrices: construct validity». *Journal of psychology*, 62, 205-209.
- BOLTON, F. B. (1955): «Experiments with the Raven's Progressive matrices, 1938». *Journal of educational research*, 48, 629-633.
- GÓMEZ FERNÁNDEZ, D. (1982): «Propiedades psicométricas del P.M.C.— Raven a partir de respuestas emitidas por escolares gallegos de E.G.B.» *Rev. de Psicología General y aplicada*, 113, V. XXVI, 777-790.
- RAVEN, J. C. (1939): «Experiments on physically and Mentally defective children with Perceptual Test». *Brit.J. Med. Psychol*; 18, 40-43 (part. I).
- RAVEN, J. C. (1948): «The comparative assessment of intellectual ability». *Brit. J. Psych*, 39, 12-19.
- RAVEN, J. C. (1948): *Progressive Matrices (1947)*, Sets I & II. Plan and Use of the scale as an individual and as a group test. Obtainable from J. C. Raven.
- RAVEN, J. C. (1970): *P.M.-56 Matrices Progresivas Escala general*. Madrid. MEPSA: Galaxia.

UN PRIMER ACERCAMIENTO A LAS ESTRATEGIAS DE RESOLUCIÓN EN LA TAREA DE CUBOS

N. Orellana, J. M. Suárez, R. Bo y F. Aliaga
Universitat de València*

La tarea de cubos, propuesta inicialmente por KOHS (1923) y utilizada y adaptada en numerosas pruebas de diagnóstico individual (Grace Arthur Scale, Goldstein-Schierer Cube Test, Grassi Block Substitution Test, Satz Block Design Rotation Test, las Escalas Wechsler, entre otras), sigue teniendo hoy en día un gran interés. A mitad de los años setenta, el estudio de esta tarea se orientó desde una perspectiva cognitiva, trabajos como los de Royer (1977), Royer y Weitzel (1977) y Schorr, Bower y Kiernan (1982) se centran en la determinación de variables cognitivas —Cohesión Perceptiva, Incertidumbre de la Tarea, Tamaño de la Serie Equivalente, Pistas de Lado, etc.— mientras que otros estudios se centran en la determinación de estrategias de solución empleadas por los sujetos (Jones y Torgesen, 1981; Beuscart-Zephir y Beuscart, 1988; Grote y Salmon, 1986; Rozencwajg, 1991). El presente trabajo se centra en esta segunda línea apuntada, pero sin perder de vista la primera ya que los diseños utilizados han sido creados en función de estas variables cognitivas.

Partiendo de anteriores trabajos y formando parte de un proyecto de investigación mucho más amplio pretendemos en este trabajo realizar un estudio sobre las diferentes estrategias de resolución en la tarea de cubos. En los trabajos realizados por otros autores sobre este tema (Jones y Torgesen, 1981; Beuscart-Zephir y Beuscart, 1988; Grote y Salmon, 1986; Rozencwajg, 1991) siempre se hace referencia a las estrategias analíticas frente a las sintéticas. Nosotros no partimos de esta dicotomía, sino que pretendemos ver cuales son las diferentes estrategias que siguen los sujetos en la resolución de los ítems. En trabajos anteriores (Jones y Torgesen, 1981; Beuscart-Zephir y Beuscart, 1988; Grote y Salmon, 1986; Rozencwajg, 1991) se ha visto que los sujetos tienden a cambiar de estrategia en función de la dificultad de los diseños más que mantener una estrategia fija (analítica/sintética). Por su parte, Rozencwajg (1991) señala un tercer tipo de sujetos, los «sincréticos», que son aquellos que identifican una figura (rombo, triángulo, trapecio, etc.) formada por diferentes cubos y a partir de ella realizan el diseño. Con ello se compara una trilogía de tipos hasta ahora propuesta y con cierto nivel de respaldo empírico, en estudios con pequeñas muestras y distintos contextos, que precisa una verificación.

* Avda. Blasco Ibáñez, 21, 46010 Valencia. Teléfono 3864430, centralita 3864420, ext. 6245.

SUJETOS

El grupo esta formado por 31 sujetos de 11 años, 15 hombres y 16 mujeres, que obtuvieron en la subprueba de Cubos del WISC-R una puntuación media de 31.52 (STD=12.5). Los sujetos no presentan características socio-demográficas especiales que los diferencie potencialmente de una población media de la Comunidad Valenciana. No obstante, la determinación del grupo no presenta elementos que garanticen inequívocamente la generalización de los resultados a esta población, tampoco era esta la finalidad central de la experiencia.

MÉTODO

A todos los sujetos se les paso la subprueba de Cubos del WISC-R de forma estándar y posteriormente una prueba experimental de cubos formada por 9 diseños de 4 cubos y 16 de 9 cubos. Estos pases se grabaron en video y posteriormente se codificaron tomando tiempos para cada cubo y las diferentes acciones que el sujeto realizaba con él (coger, rodar, girar, poner, quitar, deshacer, etc.), al mismo tiempo se dibujaba en qué lugar y en qué posición dejaba cada cubo el sujeto. De esta forma tenemos la suficiente información como para analizar las diferentes estrategias de ejecución de los sujetos. Con el fin de verlo de una forma más gráfica y reducida, se ha resumido parte de esta información en una hoja en la que se va viendo la evolución de la construcción por el sujeto (no la presentamos por razones de espacio). Para ello hemos tomado un grupo de sujetos de 11 años y hemos analizado 4 de los 16 diseños experimentales de 9 cubos. Hemos elegido estos 4 diseños (ítems 11, 15, 19 y 23) —ver figura 1— porque tienen una Cohesión Perceptiva (Royer y Weitzel, 1977) de 4, es decir, sólo 4 de los bordes internos de los cubos son del mismo color, también hemos mantenido fijo el Tamaño de la Serie Equivalente (Royer, 1977) (número de diseños distintos que aparecerían si rotáramos sucesivamente la figura 90º, la reflejáramos, o ambas cosas a la vez) en 4. Es la variable cognitiva de la Incertidumbre de la Tarea (Royer, 1977; Royer, 1984) la que varía entre 9 (todos los cubos sólidos), 12 (6 cubos sólidos y 3 diagonales), 15 (3 cubos sólidos y 6 diagonales), y 18 (todos los cubos diagonales) — ver figura 1—. Esta variable cognitiva es la que más aporta al aumento de la dificultad de la tarea (Royer, Gilmore y Gruhn, 1984; Orellana 1990, Orellana y col. 1990; Orellana y col. 1992a; Orellana y col. 1992b). Y por otra parte, estos diseños tienen un grado de dificultad intermedio para la edad de 11 años, con lo que se facilita su interpretación.

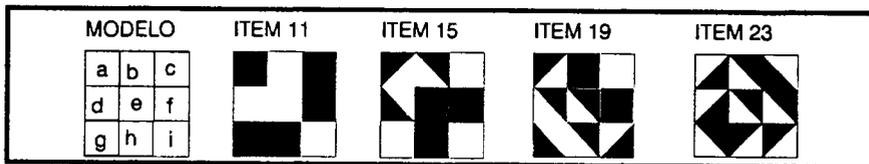


Figura 1. Muestra de los cuatro diseños utilizados en el trabajo y del modelo.

Se han tomado medidas de la puntuación directa en la subprueba de Cubos del WISC-R (Wechsler, 1974), del tiempo de realización del ítem, el número de cubos correctos, el número de pasos y tras ver y codificar las distintas ejecuciones de los sujetos, entresacamos 7 de las diferentes estrategias utilizadas, estas son:

- 1-realiza primero el cuadrado superior izquierdo.
- 2-realiza primero el cuadrado superior derecho.
- 3-realiza primero el cuadrado inferior derecho.

4-lo realiza por filas.

5-lo realiza en el sentido de las agujas del reloj.

6-realización específica en función del diseño, esto es:

-Ítem 11: realiza primero los dos rectángulos rojos (c-f o g-h),y el cubo blanco (i) y luego realiza el resto.

-Ítem 15: realizan la primera fila (a,b,c) o la primera columna (a,d,g), luego el cuadrado inferior derecho (e,f,h,i) y luego los dos cubos restantes.

-Ítem 19: realizan la primera fila (a,b,c) y luego realizan el resto.

-Ítem 23: realizan la primera fila (a,b,c) o la primera columna (a,d,g) y luego el resto.

7-otras, aquí incluimos todas las demás estrategias que solo se han dado 1 ó 2 veces a lo largo de todos los ítems, como por ejemplo: lo realiza por columnas, lo realiza asistemáticamente o cualquier colocación que no se pudiera categorizar en ninguna de las estrategias anteriormente mencionadas. por este motivo, a la hora de comentar los resultados, la referencia a esta categoría de estrategias se le dará un valor menor.

RESULTADOS

Respecto a las estrategias utilizadas por los sujetos para la realización de los diseños, se observa —ver tabla 1— que cada ítem tiene una o dos estrategias predominantes. En el ítem 11, el 48.9% de los sujetos realizan primero el cuadrado superior izquierdo (a,b,d,e) y luego el resto del diseño, mientras que un 19.4% siguen la estrategia «contraria», primero realizan los rectángulos rojos (c-f, g-h) y luego colocan el resto de los cubos. En el ítem 15, hay 3 cubos diagonales (a,b,d) y además existe cohesión entre ellos lo cual dificulta más la tarea. Aun así, un 25.8% de los sujetos realizan primero el cuadrado superior izquierdo (a,b,d,e) y un 29% empieza poniendo los cubos del cuadrado inferior derecho (e,f,h,i) y luego coloca los demás. En el ítem 19, seis de los nueve cubos son diagonales y no forma ninguna figura muy definida, quizá sea este el motivo por el cual el 41.9% de los sujetos lo realiza por filas y el 22.6% realizan primero la fila superior (a,b,c) y luego colocan el resto de los cubos. En el ítem 23, también hay un porcentaje considerable de sujetos (el 22.6%) que lo realizan por filas y un 28.8% de los sujetos colocan la primera fila (a,b,c) o la primera columna (a,d,g) y luego el resto de los cubos. Como se ha dicho anteriormente, el conjunto de estrategias incluidas en el grupo «otras» es amplio y variado, por esto, aunque se observen porcentajes de estos sujetos en esta categoría, no merece un comentario especial.

TABLA 1
LAS DIFERENTES ESTRATEGIAS DE RESOLUCIÓN. INDICANDO EL N Y EL PORCENTAJE
PARA CADA UNO DE LOS ÍTEMS

ESTRATEGIA	ÍTEM 11		ÍTEM 15		ÍTEM 19		ÍTEM 23	
	N	%	N	%	N	%	N	%
1.-Realiza 1º el cuadrado superior izquierdo	15	48.9	8	25.8				
2.-Realiza 1º el cuadrado superior derecho.					3	9.7	2	6.5
3.-Realiza 1º el cuadrado inferior derecho.			9	29.0				
4.-Lo realiza por filas.	2	6.4	2	6.5	13	41.9	7	22.6
5.-En el sentido de las agujas del reloj.	2	6.4	1	3.2	4	12.9	4	12.9
6.-Realización específica.	6	19.4	4	12.4	7	22.6	8	28.8
7.-Otros.	6	19.4	7	22.6	3	9.7	8	28.8

Respecto a los tiempos medios utilizados para cada ítem, en función de cada estrategia —ver tabla 2 y gráfica 1.a—, conviene hacer algunos comentarios. En el ítem 11, podemos decir que las diferentes estrategias no influyen mucho en el tiempo de ejecución, aunque hay que señalar que la estrategia más rápida sería la 1 —realiza primero el cuadrado superior izquierdo— utilizando 22.67 segundos. Esta misma estrategia también resulta la más rápida para el ítem 15. En el ítem 19 la estrategia más efectiva es la 2, realiza primero el cuadrado superior izquierdo, el cual contiene los únicos tres cubos sólidos del diseño. En el ítem 23 la estrategia más rápida es la 4 —realiza el diseño por filas—. Respecto a esta estrategia, hay que señalar que en los tres últimos ítems el tiempo de ejecución es muy semejante para los tres —ver tabla 2 y gráfica 1.a— independientemente de la dificultad del diseño.

TABLA 2
 LAS DIFERENTES ESTRATEGIAS DE RESOLUCIÓN. INDICANDO LA MEDIA (M) Y LA DESVIACIÓN ESTÁNDAR (STD) PARA CADA UNO DE LOS ÍTEMES, DEL TIEMPO DE EJECUCIÓN

ESTRATEGIA	ÍTEM 11		ÍTEM 15		ÍTEM 19		ÍTEM 23	
	M	STD	M	STD	M	STD	M	STD
1.-Realiza 1º el cuadrado superior izquierdo.	22.7	9.1	33.6	11.2				
2.-Realiza 1º el cuadrado superior derecho.					30.0	13.5	61.5	23.3
3.-Realiza 1º el cuadrado inferior derecho.			37.0	12.4				
4.-Lo realiza por filas.	25.0	1.41	58.5	10.6	57.2	46.8	56.1	17.2
5.-En el sentido de las agujas del reloj.	32.0	5.7	63.0	0.0	53.0	34.2	69.0	27.0
6.-Realización específica.	28.0	9.7	39.8	15.8	76.3	38.6	61.7	35.3
7.-Otros.	25.8	8.6	45.6	16.9	134.0	65.5	92.1	45.9
TOTAL	25.1	8.7	40.7	14.7	66.0	48.4	69.6	34.4

Los resultados del Cluster no son muy concluyentes, pero nos permite cierto agrupamiento entre los sujetos. Estos forman 4 grupos, más un quinto que esta formado por aquellos sujetos que no entrarían en ninguno de los 4 bloques. A este grupo lo llamamos 0 y coincide con las ejecuciones más lentas —ver gráfica 1.b—. El grupo 1, formado por 6 sujetos (19.4%), esta compuesto por individuos lentos y se caracterizan por realizar en su mayoría los ítems 11 y 23 con la estrategia 7 —otras— y el 19 con la 6, es decir, la específica del diseño —realizan primero la fila superior y luego el resto de los cubos—. El grupo 2, está integrado en general por los sujetos más rápidos —ver gráfica 1.b— y se caracterizan por utilizar la estrategia 1 en el ítem 15, la estrategia 4 o la 5 en el 19, es decir, lo realizan por filas o en el sentido de las agujas del reloj; y el ítem 23 la estrategia 6, es decir, la específica del diseño —realizan primero la fila superior o la primera columna y luego el resto de los cubos—. El grupo 3, se podría decir que es el intermedio, en lo que a tiempo de ejecución se refiere. Estos sujetos se caracterizan por realizar el ítem 15 colocando la primera fila o la primera columna y luego el cuadrado inferior derecho —estrategia 6—; y los ítems 19 y 23 en el sentido de las agujas del reloj. Por último, el grupo 4 es el más numeroso, formado por 15 sujetos (48.9%), se podría decir que es el grupo tipo, ya que, en su mayoría, utilizan las estrategias más comunes, así el ítem 11 lo realizan mayoritariamente siguiendo la estrategia 1 —construyendo primero el cuadrado superior izquierdo y luego el resto—; el ítem 15 lo hacen empezando por el cuadrado inferior izquierdo y posteriormente colocan los cubos restantes —estrategia 3—; y los ítems 19 y 23 por filas.

Es importante señalar, en relación a la ejecución particular de cada sujeto, que ninguno de ellos realiza más de dos de los cuatro ítems siguiendo la misma estrategia.

ESTRATEGIA	ÍTEM 11		ÍTEM 15		ÍTEM 19		ÍTEM 23	
	N	%	N	%	N	%	N	%
1.-Realiza 1º el cuadrado superior izquierdo.	15	48.9	8	25.8				
2.-Realiza 1º el cuadrado superior derecho.					3	9.7	2	6.5
3.-Realiza 1º el cuadrado inferior derecho.			9	29.0				
4.-Lo realiza por filas.	2	6.4	2	6.5	13	41.9	7	22.6
5.-En el sentido de las agujas del reloj.	2	6.4	1	3.2	4	12.9	4	12.9
6.-Realización específica.	6	19.4	4	12.4	7	22.6	8	28.8
7.-Otros.	6	19.4	7	22.6	3	9.7	8	28.8

Gráfica 1. Representación de los tiempos medios en función de: a) las diferentes estrategias en cada ítem; y b) los diferentes ítems en cada uno de los grupos formados por el Cluster.

CONCLUSIONES

La Incertidumbre de la Tarea parece condicionar diferentes estrategias de solución por parte de los sujetos. Además un mismo sujeto tiende a cambiar de estrategia de solución en función de las demandas planteadas por el problema. De hecho, no hay ningún caso en que se emplee la misma estrategia para todos los diseños, en los casos más estables esto se lleva a cabo en el 50% de los diseños. La variación de estrategia es, adicionalmente, un hecho no aleatorio puesto que cuando se produce una elección estable coincide con niveles de dificultad consecutivos —por ejemplo, se utiliza la misma estrategia para los dos primeros elementos y luego se varía para los dos siguientes—.

En parte como consecuencia de lo anterior no se puede afirmar la existencia de tipos «puros», de acuerdo con la tipología que se ha mencionado previamente —analíticos, sintéticos y sincréticos—. En la mayor parte de los casos se produce un cambio de estrategias sintéticas o sincréticas —en estos diseños es muy difícil deslindarlas inequívocamente— hacia estrategias analíticas —más o menos puras— cuando la demanda de dificultad de la tarea se torna mayor. No obstante, existe una proporción muy apreciable de casos en los que no se puede reconocer tipo alguno. De hecho solo es claramente reconocible y aún con tipos mixtos en un 61.27% de casos, por lo que es muy discutible el empleo sistemático de este concepto. En esta dirección, aunque los autores (Jones y Torgesen, 1981; Schorr, Bower y Kiernan, 1982; Kiernan, Bower y Schorr, 1984; Grote y Salmon, 1986; Beuscart-Zephir y Beuscart, 1988) señalan la estrategia analítica como la más efectiva aquí se encuentran pruebas de que sólo es cierto para los niveles superiores de Incertidumbre de la Tarea —en nuestro caso en los ítems 19 y 23—.

Finalmente, cabe señalar la bondad del procedimiento para poner a prueba hipótesis sustantivas sobre las estrategias de solución y sus tipologías. Los resultados obtenidos avalan su generalización a una gama mayor de elementos estimulares y la búsqueda de relación con otras características de los sujetos —rendimiento global en la tarea, otras variables cognitivas y diferenciales, etc.—.

BIBLIOGRAFÍA

- BEUSCART-ZEPHIR, M. C. & BEUSCART, R. (1988): «Tests de performances; une méthode d'Analyse des Stratégies de Résolution. Un Exemple; le Test de Cubes du WISC-R. *European Journal of Psychology of Education*. Vol. III (1), 33-35.
- GROTE, C. & SALMON, P. (1986): «Spatial complexity and hand usage on the Block Desing Test». *Perceptual and Motor Skills*, 62, 59-67.

- JONES, R. S. & TORGERSEN, J. K. (1981): «Analysis of behaviors involved in performance of the Block Desing Subtest of the WISC-R». *Intelligence*, oct —dec. Vol. 5(4) 321-328.
- KIERNAN, R. J.; BOWER, G. H. & SCHORR, D. (1984): «Stimulus variables in the block desing task revisited: A reply to Royer». *Journal of Consulting and Clinical Psychology*, 52, 4, 705-707.
- KOHS, S. C. (1923): *Intelligence Measurement*. New York: McMillan.
- ORELLANA, N. (1990): *La Tarea Psicométrica de Cubos en la Escala WISC-R: Impacto de algunas Variables Cognitivas y Diferenciales*. Tesis de Licenciatura. Facultad de Filosofía y CC. de la Educación, Universitat de València.
- ORELLANA, N., SUÁREZ, J. M., JORNET, J. M. y SÁEZ, A. (1990): «Impacto de algunas de las variables cognitivas en la estructura psicométrica de Cubos en el WISC-R». *Revista de Investigación Educativa*, Vol. 8, 16, 661-669.
- ORELLANA, N., SUÁREZ, J. M., JORNET, J. M. y BO, R. (1992): «Influencia de variables cognitivas y diferenciales en la tarea de Cubos» *BORDON*, en prensa.
- ORELLANA, N., SUÁREZ, J. M., JORNET, J. M. y BO, R. (1992): «Las construcciones parciales en la tarea de Cubos y su aprovechamiento métrico basado en un acercamiento cognitivo» *BORDON*, en prensa.
- ROYER, F. L. (1977): «Information processing in the block desing test». *Intelligence*, 1, 32-50.
- ROYER, F. L. (1984): «Stimulus variables in the block desing task: A commentary on Schorr, Bower and Kierman». *Journal of Consulting and Clinical Psychology*, 52, 4, 700-704.
- ROYER, F. L. & WEITZEL, K. E. (1977): «Effect of perceptual cohesiveness on pattern recoding in the block desing test». *Perception and Psychophysics*, 21, 39-46.
- ROYER, F. L.; GILMORE, G. C. & GRUHN, J. J. (1984): «Stimulus parameters that produce age differences in block desing performance». *Journal of Clinical Psychology*, 40, 1.474-1.485.
- ROZENCWAJG, P. (1991): «Analysis of Problem Solving Strategies on the Kohs Block Design Test». *European Journal of Psychology of Education*, Vol. VI, nº 1, 73-88.
- SCHORR, D., BOWER, G. H. & KIERMAN, R. (1982): «Stimulus variables in the block desing task». *Journal of Consulting and Clinical Psychology*, 50, 479-487.
- WECHSLER, D. (1974): *Manual for the Wechsler Intelligence Scale for Children-Revised*. New York: The Psychological Corporation.

APROXIMACIÓN AL MAPA DE DIFICULTAD DEL LABERINTO 6 DE LA ESCALA WISC-R

R. M. Bo, J. M. Suárez, N. Orellana y A. Sáez
Universitat de València*

INTRODUCCIÓN

La tarea de Laberintos, originalmente elaborada por PORTEUS (1965), está muy extendida en su utilización tanto en pruebas diagnósticas como en actividades orientadas a la formación y/o recuperación de los sujetos, especialmente en la edad preescolar. No obstante, no existe ninguna teoría, entre los diversos acercamientos realizados, que encare la explicación de la tarea en sí y justifique los diversos usos y formatos de presentación con que se emplea (BO, 1992).

Cualquier tentativa, pues, por profundizar en el conocimiento de la estructura interna de la tarea aportará datos válidos tanto para su utilización métrica como para la ulterior elaboración de una teoría explicativa de la misma. En este sentido, entendemos de especial utilidad abordar el establecimiento de un conocimiento topográfico de los componentes de dificultad que están inmersos en cualquier laberinto. Para ello, consideramos de especial importancia el análisis del mismo en base a las decisiones que el sujeto tiene que tomar a lo largo de su recorrido. De esta forma, el laberinto se podría expresar en términos de algunos puntos críticos en los que el sujeto tiene que optar entre un número determinado de alternativas de acción, evaluando las consecuencias que conlleva cada una de ellas.

Por otro lado, las alternativas que configuran cada decisión tampoco son homogéneas entre sí de forma que su evaluabilidad, impacto directo sobre la dificultad del sujeto, puede ser directa o muy costosa. En este sentido se pueden establecer algunos indicadores como la longitud total del recorrido posible o el número de cambios de dirección —esquinas que debe doblar— que debe establecer el sujeto para establecer un acercamiento más riguroso a la dificultad inherente a cada alternativa que se le ofrece al sujeto.

Después de realizar este mapa analítico deberemos enfrentarlo con las ejecuciones concretas de los sujetos para determinar las características de cada uno de los puntos de decisión y sus alternativas. A tal efecto, habrá que tomar en consideración la mayor parte de las variables que ya han demostrado su utilidad en estudios anteriores (Sáez et al., 1990; BO, 1992), como los diversos tiempos de ejecución, errores, etc. y añadir las características que hemos mencionado sobre la dificultad de las alternativas.

Finalmente, parece importante seguir considerando el impacto de las variables moduladoras más importantes: edad de los sujetos y nivel alcanzado en el conjunto de la prueba. Como ya se ha

* Avda. Blasco Ibáñez, 21, 46010 - VALENCIA. Teléfono 3864430 Centralita 3864420, ext. 6245.

apreciado reiteradamente estas dos dimensiones son muy poderosas y pueden afectar de forma decisiva a la estructura de dificultad de la tarea.

Con este objetivo se ha seleccionado el elemento 6 de la subprueba de Laberintos de la escala WISC-R (WECHSLER, 1974) por que reúne algunas propiedades idóneas para la comprobación de esta estrategia de actuación. En conjunto, constituye un elemento con suficiente dificultad, contrariamente a los 4 primeros, como para tener una expresión razonable de la ejecución a través de todas las edades en consideración. Además, su complejidad alcanza el nivel suficiente como para que se pueda abordar su estudio en un trabajo de estas dimensiones, lo que no ocurre con los últimos elementos. En cualquier caso, en este elemento ya se encuentra una muestra suficientemente amplia de los tipos de decisión y de alternativa que son las claves del modelo de análisis de dificultad que proponemos en esta tarea de laberintos.

PROCEDIMIENTO

Se ha grabado en video, durante la administración de la subprueba de laberintos comprendida en la Escala de Inteligencia WISC-R, a una selección de 73 sujetos con edades de 7, 11 y 15 años. De este modo se puede observar el proceso que cada sujeto ha seguido para llegar o no a la resolución de los laberintos. A tal efecto, y para hacer un estudio sistemático en todos los casos, se han construido unas normas de codificación en cuanto a la información que se considera relevante.

Como ya se ha mencionado anteriormente, la estrategia utilizada se basa en los árboles de decisión y concretamente en este trabajo hablaremos del árbol de toma de decisiones del ítem 6. Dicho árbol comprende 8 decisiones (D1, D2, D3, D4, D5, D6, D7, D8), como se puede ver en el gráfico 1, y cada una de ellas tiene dos caminos posibles a los que llamamos alternativas (A1 y A2). En todas las decisiones la alternativa correcta es la que se ha numerado con 2, siendo la alternativa numerada como 1 la incorrecta. Dicha incorrección consiste en meterse en una calle sin salida, salvo en la decisión 1 en la cual ambas alternativas son incorrectas.

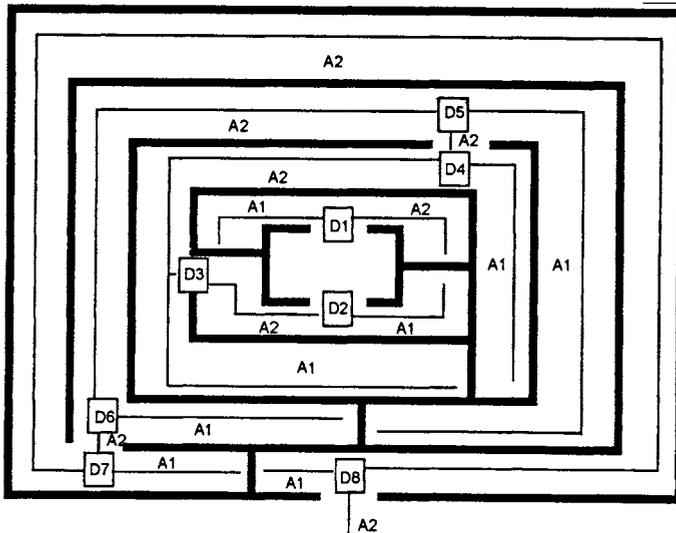


Gráfico 1. Representación de las decisiones (D1 hasta D8) y alternativas de las mismas (A1 y A2) en la solución del laberinto 6 de la escala WISC-R.

Las variables que hemos considerado relevantes, y por lo tanto codificado, son las siguientes:

* El tiempo que se ha tardado en realizar el trazado comprendido entre 2 puntos de decisión sin ninguna interrupción, y se ha codificado como TD1A1, TD1A2, TD2A1, TD2A2, TD3A1, TD3A2, TD4A1, TD4A2, TD5A1, TD5A2, TD6A1, TD6A2, TD7A1, TD7A2, TD8A1 y TD8A2, respectivamente.

* El número de esquinas que hay que doblar en cada uno de los tramos, ha sido codificado como sigue: NED1A1, NED1A2, NED2A1, NED2A2, NED3A1, NED3A2, NED4A1, NED4A2, NED5A1, NED5A2, NED6A1, NED6A2, NED7A1, NED7A2, NED8A1, NED8A2. En esta variable hay que tener en cuenta que no todos los tramos tienen el mismo número de esquinas, pudiendo oscilar entre 0 y 2 esquinas. El hecho de considerar esta variable como relevante viene dado especialmente por los tramos incorrectos, ya que pensamos que no es lo mismo meterse en una calle sin salida hasta el final, que únicamente meterse un poco y al darse cuenta salir de ahí.

* El número de veces que se ha realizado dicho tramo, ha sido codificado respectivamente como: NVD1A1, NVD1A2, NVD2A1, NVD2A2, NVD3A1, NVD3A2, NVD4A1, NVD4A2, NVD5A1, NVD5A2, NVD6A1, NVD6A2, NVD7A1, NVD7A2, NVD8A1, NVD8A2. Dicha variable al igual que la anterior adquiere mayor importancia en los tramos incorrectos, ya que no es lo mismo meterse una sola vez en un sitio incorrecto que ser reiterativo y meterse más de una vez en la misma calle sin salida.

Teniendo en cuenta los resultados obtenidos en estudios precedentes (BO, 1992) se ha optado por considerar importantes las detenciones y los errores 2 —levantar el lápiz del papel—. Ambos son indicadores de planificación forzada —efectuada sobre la marcha al ser inducida por la dificultad planteada por la tarea— y como tal han sido codificadas en las siguientes variables: NDE2D1A1, NDE2D1A2, NDE2D2A1, NDE2D2A2, NDE2D3A1, NDE2D3A2, NDE2D4A1, NDE2D4A2, NDE2D5A1, NDE2D5A2, NDE2D6A1, NDE2D6A2, NDE2D7A1, NDE2D7A2, NDE2D8A1, NDE2D8A2.

Por su parte, el tiempo que se ha empleado en las Detenciones y Errores 2 en cada tramo. Se recoge en: TDE2D1A1, TDE2D1A2, TDE2D2A1, TDE2D2A2, TDE2D3A1, TDE2D3A2, TDE2D4A1, TDE2D4A2, TDE2D5A1, TDE2D5A2, TDE2D6A1, TDE2D6A2, TDE2D7A1, TDE2D7A2, TDE2D8A1, TDE2D8A2.

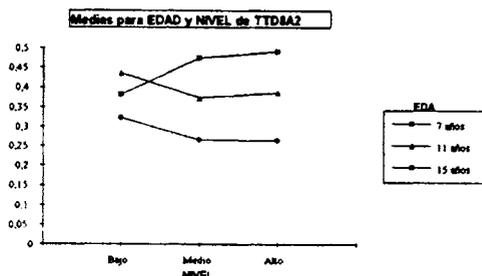
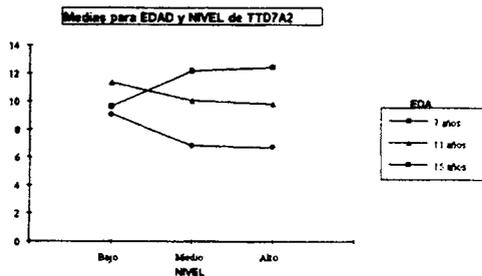
* Para finalizar se ha creado una nueva variable que es el tiempo total empleado en el tramo, que no es otro que la suma del tiempo empleado en realizar el tramo sin ninguna interrupción y el tiempo empleado en Detenciones y Errores 2 en dicho tramo. Esto se ha codificado como: TTD1A1, TTD1A2, TTD2A1, TTD2A2, TTD3A1, TTD3A2, TTD4A1, TTD4A2, TTD5A1, TTD5A2, TTD6A1, TTD6A2, TTD7A1, TTD7A2, TTD8A1 y TTD8A2.

El primer objetivo es centrar los puntos del diseño —decisiones— y los aspectos de los mismos —tiempos, repeticiones, cambios de dirección, etc.— que son cruciales para determinar su dificultad. Por ello, hemos tratado de verificar las diferencias de las variables mencionadas en los párrafos anteriores en función de la Edad y el Nivel de ejecución de los sujetos en el conjunto de la prueba de Laberintos. Las hipótesis se ponen a prueba mediante modelos de análisis de varianza mediante la rutina MANOVA del SPSS. En la tabla 1 se recogen las probabilidades correspondientes a las hipótesis que se ponen a prueba en cada caso. En este punto vamos pues a comentar únicamente los aspectos más relevantes.

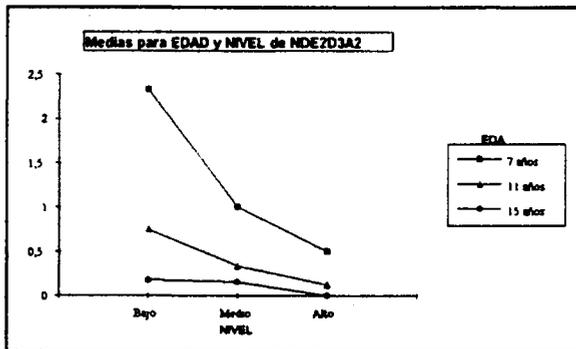
Todas las variables referentes a la primera decisión en su alternativa 2 (D1A2) resultan irrelevantes, al igual que las correspondientes a la decisión cuarta, tanto en su alternativa incorrecta como en la alternativa correcta —D4A1 y D4A2 respectivamente—. También resultan irrelevantes las variables de la sexta decisión en su alternativa correcta (D6A2) y en su alternativa incorrecta (D6A1), y las variables de la séptima y octava decisiones tanto en su alternativa correcta como incorrecta —D7A1, D7A2, D8A1, D8A2—. En todos estos tramos se puede dar algún caso aislado, pero normalmente las

TABLA 1
PROBABILIDADES DE LAS F EN LOS CONTRASTES DE HIPÓTESIS PARA LOS EFECTOS DE
EDAD, NIVEL Y SU INTERACCIÓN SOBRE LAS DIVERSAS VARIABLES DEPENDIENTES DE
EJECUCIÓN EN LA TAREA.

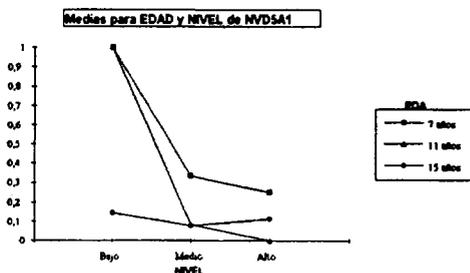
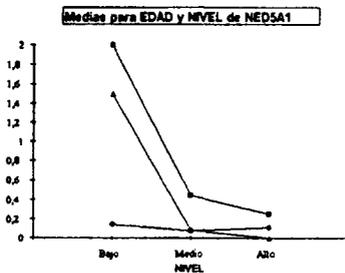
RESULTADOS							
VARIABLE	EDAD	NIVEL	EDAD x NIVEL	VARIABLE	EDAD	NIVEL	EDAD x NIVEL
TD1A1	.005	.000	.010	NED5A2	.002	.004	.001
NED1A1	.047	.014	.193	NVD5A2	.002	.004	.001
NVD1A1	.011	.000	.012	TD2A2	.007	.975	.785
TDE2D1A1	.015	.047	.004	TD3A2	.007	.975	.785
TD2A1	.002	.004	.001	TD4A2	.007	.975	.785
NED2A1	.002	.004	.001	TD5A2	.007	.975	.785
NVD2A1	.002	.004	.001	TD6A2	.007	.975	.785
TDE2D2A1	.002	.004	.001	TD7A2	.007	.975	.785
NVD2A2	.023	.683	.849	TD8A2	.007	.975	.785
TDE2D2A2	.001	.024	.006	TTD1A1	.015	.002	.010
NED3A1	.017	.448	.826	TTD2A1	.002	.004	.001
NVD3A1	.019	.370	.649	TTD2A2	.001	.027	.010
NVD3A2	.003	.653	.815	TTD3A2	.006	.100	.977
TDE2D3A2	.025	.038	.790	TTD4A1	.023	.055	.021
TDE2D4A1	.023	.055	.021	TTD4A2	.008	.007	.005
TDE2D4A2	.012	.007	.004	TTD5A1	.002	.001	.032
TD5A1	.000	.000	.007	TTD7A2	.012	.965	.674
NED5A1	.000	.000	.000	TTD8A2	.007	.975	.785
NVD5A1	.003	.000	.016	NDE2D3A2	.002	.003	.570



Gráficas I y II. Medias para el tiempo total empleado en los tramos correspondientes a las D7 y D8 en su alternativa correcta.



Gráfica III. Medias para el número de Detenciones y Errores 2 que se han cometido en la D3 en su alternativa correcta.



Gráficas IV y V. Medias del número de esquinas y número de veces que han realizado el tramo correspondiente a la D5 en su alternativa incorrecta.

alternativas incorrectas (A1) no las ha realizado ningún sujeto y las alternativas correctas (A2) todos los sujetos las realizan una sola vez y con la totalidad de cambios de direcciones posibles. Sólo hay un sujeto de la edad 7 años y nivel bajo en la resolución de la tarea global de laberintos que abandona el laberinto en la quinta decisión correcta (D5A2), y por lo tanto en las decisiones que siguen se va arrastrando dicho abandono. Las variables referentes a tiempo sí que resultan significativas por edades y por nivel, pero eso se debe a que dependen de la velocidad con la que se realiza el tramo, lo cual refleja otro componente de la dificultad del laberinto. Esto se puede ver en las gráficas I y II.

La primera decisión en su alternativa incorrecta (D1A1) muestra algún sujeto, en las edades de 7 y 11 años, que comete tal error. Concretamente son 2 sujetos de 7 años pertenecientes a nivel bajo y medio en ejecución de la tarea global de laberintos y 2 sujetos de 11 años pertenecientes a nivel bajo. No obstante, los sujetos de 11 años emplean menor tiempo en la rectificación de este error.

La segunda decisión en su alternativa incorrecta (D2A1) sólo se da en la edad de 7 años con nivel bajo en resolución de la tarea, lo que supone una gran consistencia respecto a la dificultad de la misma.

Al realizar la alternativa correcta de la segunda decisión (D2A2) es la edad de 7 años, en todos sus niveles de resolución de laberintos, los que pasan más de una vez por ella. Los sujetos de 11 y 15 años sólo la realizan una vez. Esto es debido a que los sujetos de 7 años, al encontrar problemas en la D3, tienden a solucionarlos reconsiderando las decisiones anteriores.

En la tercera decisión la alternativa incorrecta (D3A1) es elegida por sujetos de las edades de 7 y 11 años —nunca en la de 15 años—. No se dan diferencias en estas edades por la frecuencia con que se encuentra tal error sino en cuanto a la gravedad del mismo —profundidad o nº de cambios de dirección en la orientación equivocada—, siendo mayor cuanto menor es el nivel de los sujetos en la tarea.

La alternativa correcta de la D3 (D3A2), en los sujetos de 7 años, se realiza repitiendo más de una vez su recorrido. Además, hay que añadir que va aumentando el nº de veces que realizan el recorrido a medida que pertenecen a un nivel más bajo en la resolución de laberintos. Esto se debe a las mismas cuestiones comentadas respecto a la segunda decisión en su alternativa correcta. En este tramo, los sujetos parecen percibir la dificultad del problema, con lo que se ven forzados a detenerse, supuestamente para planificar la solución del problema, como se puede ver en la gráfica III.

La decisión quinta en su alternativa incorrecta (D5A1) la toman prácticamente los sujetos de todas las edades, excepto los de 11 años en nivel alto de resolución de la tarea. En las gráficas IV y V se puede ver como los sujetos de 7 y 11 años con nivel bajo cometen todos el error una vez, a diferencia de los de 15 años en los que es mucho menos frecuente. Sin embargo, en este mismo nivel, la profundidad del error no es la misma para los de 7 años que para los de 11, ya que los primeros todos se introducen hasta el fondo de la calle sin salida, y los de 11 años tienden a no llegar hasta el final. Esta misma explicación se puede dar en los otros niveles, aunque como se puede ver la diferencia entre los niveles es bastante fuerte dentro de una misma edad, sobre todo en las edades de 7 y 11 años, siendo un poco más suave en la edad de 15 años. En esta última edad y en su nivel alto parece haber algún sujeto anómalo, que es el responsable de la modificación de la posición relativa de este grupo de edad respecto a los otros.

La alternativa correcta de la decisión quinta (D5A2) es realizada por todos los sujetos de todas las edades —7, 11 y 15 años— sin ningún problema, excepto alguna anomalía en la edad de 7 años con un sujeto perteneciente al nivel bajo de resolución de laberintos que abandona el laberinto en dicho punto.

CONCLUSIONES

A través de los resultados expuestos en el apartado anterior se puede concluir que son la tercera decisión y la quinta, las decisiones en las que radica la dificultad del laberinto. Se ha podido observar que a partir de la quinta decisión, salvo algún sujeto que se sale de la norma —abandono de un sujeto de 7 años de nivel bajo en ejecución de la tarea y algún otro que se introduce en la sexta decisión en su alternativa incorrecta—, todos los sujetos resuelven perfectamente el laberinto sin meterse en ninguna calle sin salida. Así, una vez elegida la solución correcta para la quinta decisión puede decirse con rotundidad que está resuelto el laberinto.

En todas las edades la dificultad no radica en el mismo punto. Como hemos podido ver en la tercera decisión son los sujetos de 7 y 11 años los que presentan problemas, y escogen la alternativa incorrecta. Sin embargo, a pesar de no haber diferencias entre estas dos edades, sí que las hay dentro de cada una de ellas según pertenezca a un nivel u otro de ejecución. En cambio, para los sujetos de 15 años dicha decisión no plantea ningún problema. También hay que destacar en favor de los sujetos de 11 años frente a los de 7 años que mientras los primeros al cometer el error, salen y siguen hacia adelante; los segundos —7 años— vuelven a revisar las decisiones anteriormente tomadas —vuelven atrás—. Se observa, al mismo tiempo, que todos los sujetos de todas las edades perciben la dificultad del problema antes de llegar a la cuarta decisión y se toman su tiempo para planificar la solución —presentan un determinado número de Detenciones y Errores 2—. Aunque en función de la edad, varía la solución que le dan al problema en el sentido que acabamos de comentar.

En la quinta decisión tiene un gran atractivo la opción incorrecta (D5A1), de hecho la práctica

totalidad de los sujetos toman esta opción. En este caso, las diferencias que se aprecian en función de la edad y el nivel se reflejan en la «profundidad» del error, es decir, el número de cambios de dirección que realiza el sujeto en la orientación equivocada antes de percatarse de ello.

Finalmente, señalar la importancia en todos los casos de las medidas de tiempo para reflejar diferencias por edad y nivel en la tarea. Esto es consistente con resultados precedentes (B0, 1992) y refuerza la necesidad de la incorporación de esta dimensión en la valoración psicométrica de la ejecución de la tarea.

BIBLIOGRAFÍA

- BO, R. M. (1992): *Validación de la tarea psicométrica de Laberintos*. Tesis de Licenciatura no publicada. Universidad de Valencia.
- PORTEUS, S. D. (1965): *Porteus Maze Test: Fifty years application*. Palo Alto. California: Pacific Books.
- SÁEZ, A., SUÁREZ, J. M., JORNET, J. M. y ORELLANA, N. (1990): «Algunos componentes cognitivos en la resolución de la tarea de laberintos de la escala WISC-R». *Revista Investigación Educativa* —Vol. 8— nº 16, 651-659.
- WECHSLER, D. (1974): *Manual for the Wechsler Intelligence Scale for Children-Revised*. New York: The Psychological Corporation.

ALGUNOS DATOS SOBRE LA ESTRUCTURA DE DIFICULTAD DEL ELEMENTO CABALLO DE LA SUBPRUEBA ROMPECABEZAS DEL WISC-R

L. Salavert, J. M. Suárez, C. Belloch y N. Orellana
Universitat de València*

La tarea de Rompecabezas es algo muy habitual en educación tanto en las pruebas diagnósticas como en diversos materiales que tienen por objeto el desarrollo y/o la recuperación de los sujetos. En este sentido, la tarea de Rompecabezas forma parte de diferentes escalas que miden habilidades cognoscitivas, como Merrill-Palmer Scale of Mental Tests, Performance Tests of Intelligence, Arthur Point Scale of Performance Tests, diversas escalas de Wechsler, entre otras.

Por otra parte, esta tarea forma parte cotidianamente de los conjuntos de actividades de los diversos programas de entrenamiento, especialmente los orientados a la maduración en las etapas previas a la adquisición de las habilidades lecto-escritoras.

A pesar de ser, por tanto, una tarea habitual para los educadores, psicopedagogos y niños, los conocimientos que tenemos sobre la misma son muy limitados y su utilización está más basada en la intuición y acumulación de información cualitativa que en un enfoque riguroso y teóricamente bien fundado. En este sentido, parece pues muy razonable tratar de ahondar en los diversos pormenores de la tarea para ganar conocimiento sobre la misma que nos permita posteriormente llevar a cabo las primeras construcciones teóricas.

Así, la mayor parte de los estudios realizados sobre la tarea son o totalmente teóricos (Meeker, 1975), o basados en intuiciones derivadas de la actividad diagnóstica (Bannatyne, 1979; Stewart et al., 1973), o preocupados por las construcciones teóricas derivadas de la estructura factorial de las pruebas (Kaufman, 1975; Martín et al., 1977; Snow, 1980).

En este sentido, nos planteamos ahondar en la estructura factorial de esta tarea como primer objetivo, completando los trabajos previos que ya han dado algunos pasos en esta dirección (Suárez, 1987; Suárez et al., 1990). Para ello, trataremos de ver las características concretas de la actuación del sujeto con las diversas piezas que componen el rompecabezas, de forma que nos permita inducir la importancia diferencial de las mismas en la construcción del modelo mental que el sujeto precisa para la reconstrucción. Dada la importancia de las variables moduladoras críticas como la edad de los sujetos o el nivel general de su habilidad, trataremos de ver cual es su influjo en la estructura de dificultad resultante.

* Avda. Blasco Ibáñez, 21, 46010 Valencia. Teléfono directo 3864430, centralita 3864420, ext. 6245.

Finalmente, teniendo presente la importancia de relacionar los hallazgos que se planteen en cuanto a la estructura diferencial con las consecuencias métricas para la tarea tal y como está formulada, nos acercaremos a una primera valoración del aporte de esta nueva información a la variable métrica de la prueba.

Para todo este proceso, hemos elegido el elemento CABALLO perteneciente a la subprueba Rompecabezas de la escala WISC-R (Wechsler, 1974). Este elemento reúne, aparte de otras consideraciones métricas, unas propiedades que lo hacen diferente de los demás en sentido de que la estructura cognoscitiva de las piezas está más uniformada que en otros casos (Thomas, 1984). Así, en este caso las pistas cognitivas de todas las piezas se derivan de la forma de las mismas, eliminando las informaciones complementarias —dibujos, superpuestos— que aparecen en los restantes rompecabezas. Siendo que, además, las propiedades métricas de este elemento (Suárez, 1987) y de algunos aspectos de su realización (Suárez et al., 1990) son especialmente atractivos, todo ello lo convierte en la representación idónea de la tarea sobre la que efectuar la prueba de estrategias apuntada.

PROCEDIMIENTO

Se ha procedido a grabar en vídeo la ejecución de 89 sujetos en la subprueba de Rompecabezas del WISC-R (WECHSLER, 1974). Por lo que respecta a las medidas sobre la actuación del sujeto en la tarea, se han codificado las acciones que puede llevar a cabo sobre cada pieza componente. Estas acciones se han sintetizado en seis tipos básicos:

— **Mirar**: reúne todas las acciones en las que el sujeto dirige claramente su atención a una pieza en concreto, sin que intervenga ningún otro tipo de actuación.

— **Desplazar**: reúne todas las acciones en las que el sujeto toma una pieza y la desplaza. Normalmente se producen acciones de este tipo cuando el sujeto está considerando posibilidades de acoplamiento con otras piezas. Se refleja aquí, únicamente, cuando no se producen uniones mínimamente estables con ninguna otra pieza.

— **Acercar**: reúne todas las acciones en las que se produce una unión provisional de dos o más piezas. El carácter de provisionalidad de la unión lo otorga, obviamente, el devenir el conjunto del proceso de resolución de la tarea. Así, una determinada unión no se puede determinar como definitiva o provisional hasta que no se contempla todo el desarrollo de un sujeto sobre el problema.

— **Quitar**: reúne todas las acciones en las que tras un acercamiento o unión provisional con otra pieza se rompe el acoplamiento y se retira la pieza que estaba ensamblada.

— **Girar**: reúne todas las acciones sobre cambios de orientación o rotaciones que se imprimen a una pieza que el sujeto ha tomado previamente en la mano. Normalmente, estos giros se tienden a producir en las piezas cuya unión es más ambigua, cuyas pistas cognitivas son menos claras.

— **Unir**: reúne todas las acciones encaminadas al acoplamiento con carácter definitivo de las piezas. Como se acaba de señalar en las acciones de acercamiento, su discriminación sólo se puede producir con la visión conjunta del proceso de resolución del rompecabezas por el sujeto.

Estas acciones se han reflejado para cada una de las piezas componentes, que en el caso del CABALLO son: la cabeza (codificada CABCA), el tronco (CABTRO), el lomo (CABLO), la pieza con el rabo (CABRA), la pata delantera (CABPD) y la pata trasera (CABPT). Las variables definitivas se producen por la combinación de las seis acciones con las seis piezas, situándose la inicial de la acción al final del nombre codificado. Así, para la cabeza del caballo se tienen las siguientes variables: CABCA_M, CABCA_D, CABCA_Q, CABCA_A, CABCA_G y CABCA_U; para el tronco: CABTRO_M, CABTRO_D, CABTRO_Q, CABTRO_A, CABTRO_G y CABTRO_U; y según este mismo procedimiento para las restantes piezas.

RESULTADOS

En primer lugar, nos centraremos en el objetivo de profundizar en la estructura de dificultad interna al rompecabezas. Esto lo haremos analizando la atención diferencial que los sujetos otorgan a las diversas piezas y los diversos tipos de acciones que realizan en cada caso. Para ello, es importante determinar si existen diferencias entre las diversas piezas en cada uno de los tipos de acciones seleccionado y, además, si estas diferencias son iguales a través de los tres niveles de edad estudiados —7, 11 y 15 años—. Para ello, se establecen modelos ANOVA de tipo mixto, con las variables independientes EDAD —entre— y PIEZAS —intra— y utilizando como covariada la puntuación del sujeto en la prueba —NIVEL—. La variable NIVEL es tan importante, como se ha demostrado en ocasiones precedentes (Suárez, 1987; BO, 1992), que no se pueden extraer consecuencias diferenciales a través de la EDAD sin tomarla en consideración.

El patrón resultante de promedios de acción con las diversas piezas del Caballo muestra algunos puntos de especial interés —ver tabla 1 y fig. 1—:

— Se encuentra una estructura diferencial entre las piezas que permanece esencialmente constante para casi todas las acciones. Esto se verifica cuantitativamente en las elevadas significaciones obtenidas entre las diversas piezas en todas las acciones estudiadas —ver tabla 1—. Además, cualitativamente se aprecia la existencia de este patrón constante en las representaciones gráficas que se recogen en la Figura 1. Así, la mayor parte del interés se concita en torno a las piezas del Tronco, Lomo y Rabo de la figura. De hecho, las restantes piezas aglutinan muchas menos acciones de todo tipo.

— La tríada de piezas fundamentales se configura de forma que el Tronco y el Rabo reúnen mayor número de acciones que el Lomo. Esto es una constante para la mayor parte de las acciones, con la excepción de las de Girar y Unir. Además, este patrón permanece estable con ligeras modificaciones para los dos niveles de edad más bajos, rompiéndose sistemáticamente la tendencia a los 15 años de edad. Esto puede parecer, a primera vista, sorprendente ya que las piezas del Tronco y del Rabo reúnen menos dificultad que el Lomo. No obstante, al tratarse de acciones encaminadas al ensamblaje del modelo final, es precisamente la dificultad ofrecida para «encajar» la pieza intermedia (Lomo) en el modelo mental lo que conduce a un mayor número de acciones de diverso tipo con las tres piezas. Particularmente, estas acciones se centran más en las dos piezas extremas puesto que «se ven con mayor claridad». En este sentido, conviene recordar, por ejemplo, la posibilidad de montar un caballo «reducido» suprimiendo la pieza del Lomo en su construcción.

— La acción de Girar es una de las más importantes entre las que se han segmentado, ya que tiene conexiones evidentes con los estudios de rotación mental (Cooper y Shepard, 1973; Kosslyn, 1980). Esta acción demuestra un indicio claro sobre la dificultad de «encajar» la determinada pieza en el modelo, lo que hace necesario para el sujeto reiterar las operaciones de ajuste. En este sentido, se aprecia una estructura semejante a la ya comentada para los casos anteriores, salvo lo que se refiere a las tres piezas centrales del rompecabezas. Aquí, se observa un patrón de dificultad creciente desde el Rabo, pasando por el Lomo y acabando en el Tronco. La menor dificultad del Rabo parece esperable precisamente por la presencia de este apéndice que reduce considerablemente las opciones de unión con las restantes piezas —todo ello completado por el aditamento unívoco de su unión con la pieza de la Pata Trasera—. Las dos restantes piezas, no obstante, muestran una dificultad inversa a lo esperado, sin duda por el efecto holista creado por la incertidumbre de colocación de la pieza intermedia, que conduce a una minimización de las pistas proporcionadas por la unión de la Pata Delantera o la diagonal de unión con la Cabeza. Esto se encuentra, además, respaldado por el hecho de que a los 15 años de edad se rompe la estructura comentada y los sujetos muestran una dificultad creciente para el Rabo, el Tronco y el Lomo, sucesivamente. Así, cuando los sujetos tienen una mayor madurez y/o

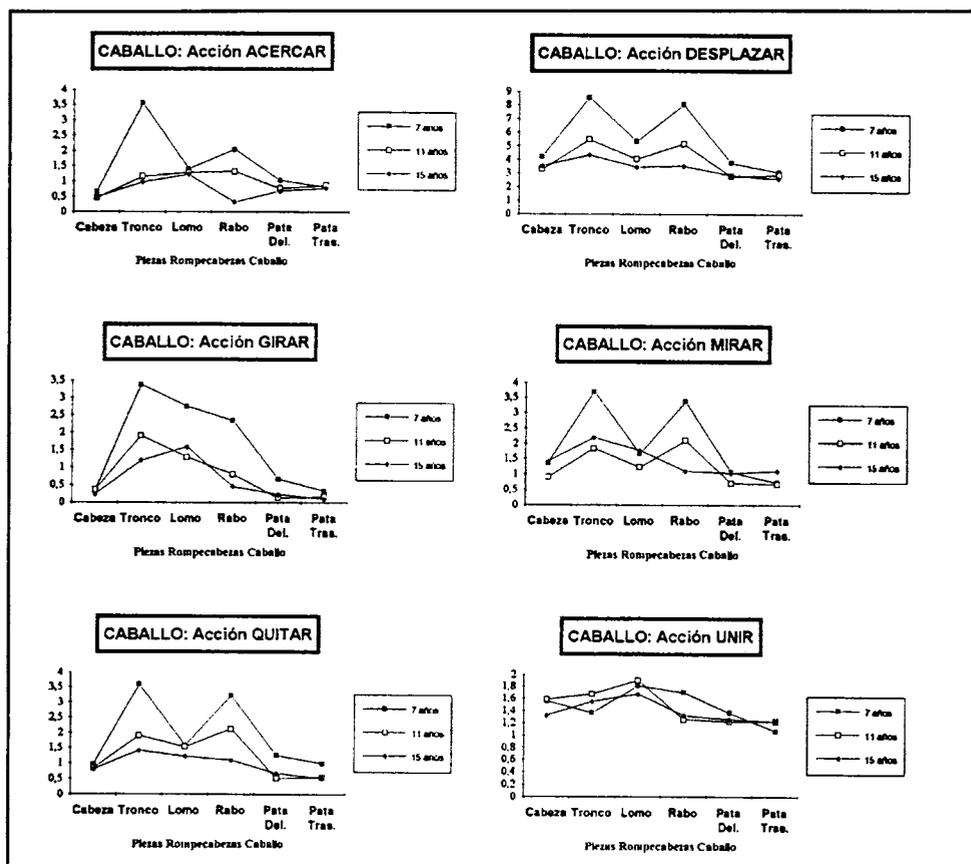


Figura 1. Representaciones gráficas de los promedios de ejecución de las diferentes acciones en cada una de las tres edades para las seis piezas que componen el caballo.

nivel de habilidad aislan perfectamente el problema y lo reducen a sus justos términos, con lo que la estructura de dificultad «emergente» se corresponde de modo estricto con el patrón anticipado.

— Las acciones, en general, muestran una correcta diferenciación para las tres edades en cuestión —7, 11 y 15 años—. De este modo, con la excepción de las acciones de Unión definitiva de las piezas, que no muestran ninguna diferencia entre las edades, en la mayor parte de los casos restantes encontramos un perfil suficientemente diferenciado entre ellas. No obstante, las diferencias no son homogéneas, sino que se acentúan entre los 7 años y las dos restantes edades. Este hecho resulta concordante con los resultados métricos y diferenciales que se han comentado en estudios previos (Suárez, 1987), ya que ha quedado bien establecida la existencia de un efecto techo notable para la tarea en torno a los 10-11 años de edad.

TABLA 1
 PROBABILIDADES CORRESPONDIENTES A LOS EFECTOS DE LOS MODELOS ANCOVA
 PARA LAS DIFERENTES ACCIONES

HIPÓTESIS

ACCION	EDAD	PIEZAS	ExP	NIVEL
MIRAR	n.s.	.000	.010	n.s.
DESPLAZAR	n.s.	.000	.004	.027
GIRAR	(.093)	.000	n.s.	.002
QUITAR	n.s.	.000	.011	.003
ACERCAR	n.s.	.000	.000	n.s.
UNIR	n.s.	.000	n.s.	n.s.

Tras haber revisado algunas de las características de la actuación de los sujetos con las diferentes piezas ya se ha apuntado una relación muy considerable con el rendimiento global del sujeto, a través de las variables edad y nivel general en la prueba. No obstante, dado que estamos manejando un conjunto de variables nuevas, parece especialmente pertinente no sólo determinar las características descriptivas y/o diferenciales de las mismas, sino también afrontar su conexión con los otros indicadores de rendimiento. Este proceso nos ayudará a comprender mejor la importancia de la información que estamos analizando para contribuir a la medición de la ejecución del sujeto en la tarea. En este sentido, hay que tener presente que la puntuación psicométrica del elemento CABALLO está determinada por dos componentes: el tiempo total empleado en su resolución —con la limitación máxima de 150 segundos— y el número de uniones correctamente realizadas entre las piezas. A pesar de esta limitación podemos determinar la importancia relativa y la eficacia global de los diversos componentes para explicar la ejecución total en el elemento, añadiendo al modelo las variables que reflejan las acciones realizadas con las piezas. Para este propósito se ha utilizado la regresión lineal —el ajuste indica que este modelo es perfectamente apropiado para el caso— del conjunto de variables de tiempo de ejecución total, número de uniones correctas entre las piezas y acciones efectuadas con las mismas, sobre la variable dependiente que recoge la puntuación total en el ítem. Dado el carácter exploratorio de la estrategia se ha utilizado el procedimiento paso a paso para la determinación de las variables más relevantes a incluir en el modelo.

En la tabla 2 se recogen los datos más significativos de la aplicación de los modelos de regresión lineal para las tres edades estudiadas (7, 11 y 15 años). No se incluye más información por motivos de espacio, aunque estimamos que esta puede ser suficiente para la aproximación que nos interesa en este punto. A partir de la revisión de las soluciones encontradas, se pueden extraer las siguientes conclusiones:

— Existe una excelente predicción de la puntuación global a partir de estos componentes, como no debe de ser menos por otra parte. El nivel de explicación conseguido en ningún caso es inferior al 90% de la variación total en la estimación más conservadora. No obstante, se aprecia una reducción de la capacidad predictiva en el nivel de 15 años que resulta concordante con los problemas básicos de medida que se dan en esta tarea en las edades superiores y que ya se han mencionado previamente.

— Se produce una estructuración diferencial muy marcada en función del nivel de edad de los sujetos. Así, mientras en los más jóvenes la predicción se asienta en una combinación de las uniones correctas y de las acciones de los sujetos, en la edad de 11 años hay una combinación de los tres componentes —uniones correctas, tiempo y acciones— con predominancia neta de las uniones correc-

tas. En la edad superior también se dan los tres componentes pero con una predominancia fuertemente marcada del tiempo de ejecución de los sujetos.

— Las uniones correctas tienen un papel muy relevante en la explicación del rendimiento global a través de todas las edades consideradas. Así, es el componente más importante en las dos primeras edades —7 y 11 años— y es el segundo componente más relevante en la edad de 15 años. Su evolución se traduce en una pérdida progresiva de importancia que se acentúa de los 11 a los 15 años, con una pérdida de contribución a la explicación de un 32% aproximadamente.

— Los tiempos globales de ejecución son un predictor muy relevante del rendimiento de los sujetos en este elemento en los dos últimos niveles de edad. Esto sucede muy especialmente en el nivel de 15 años de edad en el que se acercan a explicar las dos terceras partes de la variación observada. El papel de esta variable sigue una tendencia creciente con la edad, lo que es coherente con el hecho de que una vez dominada una tarea las diferencias básicamente se establecen en la velocidad con que se es capaz de resolverla.

— Las acciones tienen una contribución sustancial a la explicación del rendimiento en este elemento a través de las tres edades consideradas. Esta contribución, no obstante, muestra una tendencia a reducirse con la edad, pasando de algo más de una cuarta parte de las diferencias observadas, en los 7 años, a un reducido 4% en el último nivel de edad. Esta tendencia parece congruente con el hecho de que cuanto más se domina una tarea las diferencias se deben menos al número de acciones que se realizan para resolverla o a sus tipos concretos —que tienden a uniformarse y esquematizarse— y más a la rapidez con la que se ejecutan tales acciones. Así, hay un menor número de acciones de todo tipo que, además se producen en una menor cantidad de tiempo por lo que son menos relevantes para la explicación de las diferentes ejecuciones de los sujetos.

— Por lo que respecta a las diferentes contribuciones, se puede apreciar un patrón sustancialmente diferente en función de la edad. Mientras en los 7 años las diferencias parecen concentrarse en acciones referentes a dos piezas —cabeza y rabo—, en la siguiente edad la escasa contribución se generaliza a casi todas las piezas del caballo y en los 15 años parece concentrarse en la parte final del mismo —rabo y pata trasera—.

TABLA 2

SE OFRECEN LOS DATOS GLOBALES DE LA PREDICCIÓN REALIZADA SOBRE LA PUNTUACIÓN TOTAL DEL ELEMENTO EN CADA NIVEL DE EDAD (R, R², R² AJUSTADA) Y LAS VARIABLES PREDICTORAS INCLUIDAS EN CADA ECUACIÓN JUNTO CON EL VALOR DE LA PRUEBA T PARA LA SIGNIFICACIÓN DE SU COEFICIENTE BETA Y LA EXPLICACIÓN DE VARIANZA QUE SE OBTIENE CUANDO SE INCLUYEN EN LA ECUACIÓN (ENTRE PARÉNTESIS)

	7 AÑOS		11 AÑOS		15 AÑOS	
R	.982		.990		.962	
R ²	.964		.981		.926	
R ² Ajust.	.955		.976		.915	
PREDICTORES	UNCOR	15.998 (.678)	UNCOR	28.315(.580)	TITOT	-14.730(.608)
	CABCA_G	-8.089 (.821)	TITOT	-9.351(.902)	UNCOR	9.121(.873)
	CABRA_Q	-8.303(.943)	CABTRO_G	-6.455(.931)	CABRA_U	3.265(.903)
	CABCA_U	2.935(.943)	CABRA_D	2.336(.964)	CABPT_G	-2.201(.915)
	CABRA_M	2.609(.955)	CABPD_Q	3.573(.972)		
			CABLO_U	2.254(.976)		

CONCLUSIONES

A partir de esta somera profundización en algunos aspectos de la estructura de dificultad interna de la tarea de Rompecabezas utilizando para ello el peculiar ítem del CABALLO de la prueba WISC-R se aprecian indicios suficientes para animar a la generalización y ampliación de la estrategia por esta y otras vías.

En primer lugar, se puede establecer una estrategia de verificación entre los supuestos sobre la tarea, derivados del análisis y la observación ya que no existe una teoría de base al respecto, y la evidencia empírica que se obtiene. En general, hemos encontrado respaldo para la práctica totalidad de los supuestos de partida hallando, no obstante, importantes matices al respecto. En este sentido, la modificabilidad de la estructura de dificultad de la tarea con la EDAD parece un hecho especialmente importante que convendría profundizar y dar más contenido.

Por otro lado, las medidas efectuadas sobre las ACCIONES de los sujetos han mostrado una importancia sustancial por su relación con la totalidad de la ejecución en la prueba en una situación desfavorable. De este modo, se puede vislumbrar una incidencia indirecta de la información contenida en las acciones en base a un cambio en el sistema de ponderaciones de las mismas realizadas entre las piezas. Además, su contribución es diferente en función de la EDAD de los sujetos lo que sugiere un tratamiento de la variable métrica diferenciado por edades. Por ello, se deberían ajustar las ponderaciones de las acciones teniendo presentes los niveles inferiores de edad —entre los 6 y los 10 años aproximadamente— y, al mismo tiempo realizar una reforma de las bonificaciones por tiempo teniendo presentes los niveles superiores. Esta propuesta se corresponde con los intervalos de mayor impacto de ambos tipos de información sobre la ejecución total de los sujetos.

BIBLIOGRAFÍA

- BANNATYNE, A. (1979): «Spatial competence, learning disabilities, auditory-vocal deficits and a WISC-R subtest recategorization». *Journal of Clinical Child Psychology*, Fa1 Vol. 8(3) 194-200.
- BO, R. M. (1992): *Validación de la tarea psicométrica de Laberintos*. Tesis de Licenciatura no publicada. Universidad de Valencia.
- COOPER, L. A. y SHEPARD, R. N. (1973): «Chronometric studies of the rotation of mental images». En W. G. CHASE (Ed.): *Visual Information Processing*. London: Academic Press.
- KAUFMAN, A. S. (1975): «Factor analysis of the WISC-R at 11 age levels between 6 1/2 and 16 1/2 years». *Journal of Consulting and Clinical Psychology*, 43, 135-147.
- KOSSLYN, S. M. (1980): *Image and Mind*. Cambridge, Mass.: Harvard Univ. Press.
- MARTÍN, J. D., BLAIR, G. E., STOKES, E. H. y ARMSTRONG, G. (1977): «Correlation of the object assembly and block design tests of the Wechsler Adult Intelligence Scale and the Torrance Tests of Creative Thinking». *Educational and Psychological Measurement*, 37, 1.095-1.097.
- MEKER, M. N. (1975): *Glosary for SOI Analysis*. Available from SOI Institute, 214 Main St., El Segundo, California.
- SNOW, R. E. (1980): «Aptitude Processes». In R. E. SNOW, P. FEDERICO & E. F. MONTAGUE: *Aptitude, Learning and Instruction*. Vol.: *Cognitive Process Analysis of Aptitude*. Hillsdale, N.J.: LEA.
- STEWART, D. J., POWERS, J. y GOGUAUX, C. (1973): «The Wechsler in personality assessment: object assembly subtest as predictive». *J. of Consulting & Clinical Psychology*, 40, 408.
- SUÁREZ, J. M., SÁEZ, A., ORELLANA, N. y JORNET, J. M. (1990): «Construcciones en la tarea de Rompecabezas de la Escala WISC-R y revisión métrica de la misma». *Revista de Investigación Educativa*, Vols. 8-16, 671-679.

- SUÁREZ, J. M. (1987): *Estudio Psicométrico-Diferencial de la Escala WISC-R*. Tesis doctoral no publicada. Universidad de Valencia.
- THOMAS, A. (1984): «Bender scores and the horse as a distinct ítem on object assembly of the WISC». *Perceptual and Motor Skills*, 59, 103-106.
- WECHSLER, D. (1974): *Manual for the Wechsler Intelligence Scale for Children-Revised*. New York: The Psychological Coporation.

UN CUESTIONARIO PARA IDENTIFICAR NIÑOS ALTAMENTE CREATIVOS

Judith Serrano Segura
Maristas Montserrat, Lérida

La finalidad de esta Comunicación es presentar un Cuestionario elaborado por mí. Este Cuestionario se denomina CINAC («Cuestionario para Identificar Niños Altamente Creativos»). El CINAC tiene como objetivo evaluar las actitudes que integran la capacidad creativa en los niños.

El CINAC evalúa dos dimensiones de la Actitud Creativa: la Laboriosidad (LAB) y la Flexibilidad Mental (FXM). El factor LAB evalúa el **deseo** de resolver problemas de producción **convergente**, y el **deseo** de ofrecer soluciones **elaboradas** a los problemas. Por su parte, el factor FXM evalúa la **flexibilidad** mental, la **curiosidad** e intereses variados, el **deseo** de ofrecer soluciones **originales** a los problemas, y el **deseo** de resolver problemas de producción **divergente**.

Las puntuaciones del CINAC deben utilizarse para **complementar**, nunca para substituir, puntuaciones aptitudinales de **capacidad creativa**.

Los niños que puntúan alto en la escala LAB suelen puntuar también alto en pruebas de aptitud convergente que requieran **perseverancia** y **atención**, y en pruebas de aptitud divergente donde se evalúe la **elaboración**. Sus compañeros y maestros les consideran trabajadores, detallistas e inteligentes. Su autoconcepto intelectual suele ser bueno. Su rendimiento escolar oscila normalmente entre el notable y el excelente.

Por su parte, los niños que puntúan alto en la escala FXM suelen puntuar también alto en pruebas que requieran **rapidez** y **flexibilidad**. Destacan especialmente en las pruebas de aptitud divergente donde se evalúe la **originalidad** y la **fluidez**. Sus compañeros y maestros les consideran listos, imaginativos y originales. Su rendimiento escolar suele ser bueno, pero no especialmente brillante.

La prueba consta de 75 ítems a ser aplicados en una sola sesión. Las frases son leídas en voz alta por el examinador. Los niños tienen una hoja de respuestas en las que sólo pone «si» y «no» al lado del número de cada pregunta. Los niños deben tachar el «si» o el «no» según estén o no estén de acuerdo con la frase leída.

La prueba CINAC se corrige mediante plantilla. La puntuación de cada factor es el número total de coincidencias entre las marcas del niño y los círculos de la plantilla. Una vez obtenida la puntuación directa, se estandariza mediante el baremo. El baremo del CINAC trata de discriminar bien sólo la cola derecha de la distribución. Por esta razón, resulta más adecuado trabajar con unas puntuaciones típicas «desviadas» (PTD) que siguen una distribución asimétrica.

El baremo está diferenciado por edad, pero no por sexo, ya que esta variable no influye en las puntuaciones. Asimismo, cabe destacar que el ambiente socioescolar afecta las puntuaciones del

CINAC. En caso de escuelas de destacada permisividad o autoritarismo, debe tenerse en cuenta este efecto.

La historia del «Cuestionario para Identificar Niños Altamente Creativos» (CINAC) empezó en 1982. Durante los diez años transcurridos, se ha trabajado con distintas muestras: una de 671 niños de 5º de E.G.B., otra de 146 niños de 5º a 8º, otra de 230 niños de P-5 a 6º, otra de 893 niños de P-5 a 7º, y otra de 1.014 niños más, de P-5 a 6º. En total, el CINAC ha sido aplicado a 2954 niños en alguna de sus versiones preliminares o definitiva.

Las distintas muestras de análisis y tipificación tienen en común estar constituidas por niños escolarizados en Cataluña. En la muestra de baremación están representados todos los niveles socioeconómicos. En las zonas rurales, hay escuelas unitarias, públicas comarcales, y privadas concertadas (tanto religiosas como laicas). Las zonas urbanas se refieren a Barcelona y su área metropolitana, Lleida, Reus y Tarragona. En estas zonas, se encuentran escuelas públicas de clase baja y de clase media-alta, y escuelas privadas desde concertadas a escuelas de clase alta.

Durante el tiempo de elaboración del CINAC, se realizaron estudios profundos de los 435 ítems que se han probado en total. La prueba está formada por el 17,24% de las frases mejores. El análisis de los elementos asegura la validez interna del Cuestionario.

En la selección de los ítems, algunos de los objetivos a conseguir eran que los factores:

- * tuvieran medias parecidas y dispersiones posibles equivalentes,
- * tuvieran una dificultad media lo más baja posible, para facilitar la discriminación de la cola derecha de la distribución,
- * presentaran una distribución lo más asimétrica (positiva) posible,
- * tuvieran una dispersión (relativa) máxima, y
- * fueran lo más independientes posible entre sí.

En una muestra de 2.137 niños, los descriptores de los dos factores del CINAC son:

	LAB	FXM
MEDIA	16,4815	16,3255
DT	5,8945	6,7546
RANGO REAL	1-29	1-38
RANGO TEÓRICO	0-32	0-43

La dificultad media de cada factor en una muestra de 1123 niños es la siguiente:

	LAB	FXM
DIFICULT. MEDIA	.4888	.3797

La dificultad media de LAB es un poco alta, pero los ítems son muy válidos. La dificultad de los ítems de la escala oscila entre .2125 y .8058, y la mediana está en .5131, aproximadamente. La dificultad de los ítems de la escala FXM oscila entre .0821 y .6837. La mediana está en .3688, aproximadamente.

Los datos sobre la asimetría y la dispersión relativa en la muestra de 1123 niños son los siguientes:

	LAB	FXM
COEF. ASIMETRÍA	+0,0973	+0,1200
COEF. VARIACIÓN	33,5010%	40,7794%

En la muestra de 2137 niños, la correlación entre los factores es .28272. Entonces, los dos factores se pueden considerar relativamente independientes.

Para calcular la fiabilidad del CINAC, se utilizó la fórmula 20 de Kuder-Richardson. Esta fórmula es igual al coeficiente alpha cuando se trata de ítems dicotómicos. En una muestra de 1123 niños, los coeficientes de estabilidad interna del CINAC son:

	LAB	FXM
COEF. FIABILIDAD	.7594	.8136
Spq	7,3371	9,1722

En una muestra de 2137 niños, se calculó la fiabilidad mediante la fórmula 21 de Kuder-Richardson. Los resultados son:

	LAB	FXM
COEF. FIABILIDAD	.7948	.7966

En relación con los coeficientes de fiabilidad que ofrecen normalmente los cuestionarios de personalidad, los del CINAC pueden calificarse de excelentes.

Por su parte, la validez externa del Cuestionario está también correctamente documentada. El factor LAB correlaciona significativamente con al menos veintisiete criterios. De entre ellos, destacan el rendimiento escolar, pruebas objetivas de razonamiento y percepción, puntuaciones de elaboración, y denominaciones tanto de los compañeros como del maestro de trabajo, cultura y persistencia. El factor FXM correlaciona con al menos veintitres criterios. Se pueden destacar las puntuaciones de flexibilidad, originalidad, fluidez, y razonamiento. Este factor también correlaciona significativamente con el rendimiento escolar, y con denominaciones de los compañeros y del maestro referidas a la inteligencia y la originalidad.

Como conclusión, el CINAC se presenta como un instrumento de diagnóstico útil para la intervención educativa diferenciada.

ESCALA DE CREENCIAS VALORATIVAS DE LOS PADRES

Trinidad Donoso
Universidad de Barcelona

1. FUNDAMENTACIÓN DE LA ESCALA DE CREENCIAS VALORATIVAS DE LOS PADRES

Los autores que se han ocupado de trabajar el tema del valor desde diversas escuelas consideran la importancia del medio social, las relaciones interpersonales y las figuras parentales en la adquisición y el desarrollo de los valores en el ser humano.

Gilligan (1976) habla de identificación con los padres para el establecimiento de los sentimientos de culpa y de fracaso. Maier (1982), desde la misma escuela psicoanalítica, se refiere a la adquisición de una disciplina interna y aceptación de reglas sociales en los primeros años.

Las teorías del desarrollo cognitivo hablan de las interacciones entre el ambiente, en los primeros años ambiente familiar, y el sujeto para la adquisición de conductas relacionadas con el mundo valorativo. (Hersh et. al 1984) (Kurtines, Gelwirtz, 1984) (Lockwood, 1978) (Malinowsky, Smith, 1985) (Kay, 1976) (Scharf, 1978).

Desde los planteamientos de la psicología existencial y humanista las posibilidades del ambiente son las que permiten las actualización y descubrimiento de valores, (May, 1978) (Maslow, 1962), y al mismo tiempo las situaciones vividas dentro del campo relacional son las que concretarán en cada sujeto la manera de percibir y realizar los valores.

Nosotros hemos definido el valor como todo criterio que guía y orienta la acción, representando una preferencia importante para la vida de un sujeto y que ha sido elegido libremente, (Donoso, 1992).

Bartolomé (1981) hace una distinción en un intento de operativizar los valores entre:

Creencias de valor: se trata de la dimensión cognitiva del valor, su conceptualización, bien aparezca expresada verbalmente, bien se halle implícita en el sujeto que la reconoce al contacto con determinados materiales.

Valores vivenciados: son los valores que ponemos en juego en la elección de determinados cursos de acción. Hacen referencia a la vida.

La escala de creencias valorativas surge en un intento de comprobar las preferencias o intereses de valores que desearían los padres ver en sus hijos, estas preferencias por unos valores u otros tiene que reflejarse de alguna forma en su conductas y en sus actuaciones verbales de cara a sus hijos.

Al construir una escala ofreciendo referentes de valores hemos de ser conscientes de que la tipología de valores que surgirán en este caso son creencias de valor puesto que el sujeto va a contestar en función de valores asimilados cognitivamente, podríamos decir, y no de valores que ponga en juego o exprese en una situación específica.

2. CONSTRUCCIÓN DE LA ESCALA

2.1. Determinación del objetivo a medir

Hemos de contar que en una escala de preferencias valorativas el adulto estará pensando, sobre todo, en situaciones ideales de algo que «se desearía» conseguir, aunque todavía «no sea».

El objetivo por tanto de la escala es: medir los valores ideales de los padres con respecto a sus hijos.

2.2. Elaboración de la escala

Para conseguir este objetivo elaboramos unas frases que se refirieran a valores que les gustaría a los padres conseguir en sus hijos. Elegimos frases en lugar de valores aislados porque es más fácil captar un significado profundo mediante frases ya que es otro el nivel interpretativo.

La escala debería recoger una amplia muestra de valores sin circunscribirse a ninguna categoría en especial por lo que decidimos trabajar con un sistema de categorías amplio. Este sistema de categorización fue recogido de M. Bartolomé (1981) que cuenta con una construcción teórica sólida además de con un intento de validación empírica al haberlo aplicado la propia autora a muestras reducidas de sujetos.

Construimos una amplia muestra de frases relacionadas con el sistema de categorización y pasamos el sistema de categorías y las frases a jueces que nos adjudicarían cada una de las frases a la categoría que pertenecía.

2.3. Validación

Por una parte nos interesaba que fuera revisada por jueces que tuvieran criterios sobre el estudio de los valores, es decir fueran expertos, por otra también nos interesaba tener resultados que provinieran de las personas a quienes se les iba a pasar la escala, es decir padres.

Elegimos diez jueces, cinco considerados como expertos y cinco padres.

Una vez hubieron revisado los jueces las frases elegimos dos frases para cada categoría de aquellas que obtuvieron acuerdos del 100% entre los jueces. Únicamente dos frases no cumplieron este criterio de acuerdo del 100% y se quedaban en el 80%. El criterio que seguimos para decidir qué frases, de entre todas las que habían obtenido el 100%, seleccionábamos, fue escoger aquellas que eran más generales y podían ser útiles para un número más amplio de valores.

2.4. Elección del sistema de puntuación más adecuado para la escala

Las frases se agruparon de dos en dos formando pares, cada una de las 16 frases con las otras 14 restantes. De esta forma cada una de las frases de una categoría quedaba emparejada con las restantes frases de las otras categorías. (García Hoz, 1972).

Se trataba de que los padres pudieran optar por una u otra de las frases en la combinación que se daba.

La fórmula de la combinación por pares que utilizamos fue:

$$\frac{n(n-1)}{2} \text{ — N}^\circ \text{ categorías}$$

Queríamos evitar el que las frases de una misma categoría quedaran emparejadas entre sí, lo que obligaría a una elección «extraña» para el sujeto.

Este planteamiento dio en total 112 combinaciones de frases que se agruparon al azar.

El problema que surgió una vez estuvieron agrupadas todas las frases en pares fue el criterio de puntuación que se daría a los sujetos, o se les obligaba a contestar en cada par o se les dejaba libertad de que contestaran o no.

Pasamos una muestra del cuestionario a 10 padres para establecer si se incluía la elección forzosa o no y averiguar qué clase de puntuación matizaba mejor.

No sólo pasamos el cuestionario como prueba y a unos sujetos se les pidió que debían elegir obligatoriamente una frase de cada par y a otros se les permitió no elegir ninguna frase del par si así lo deseaba, también entrevistamos a los sujetos para que nos dieran sus impresiones ante cómo se habían sentido al tener que rellenar de una manera u de otra.

Las primeras conclusiones es que la elección forzada oprime el sujeto cuando estamos trabajando en el mundo de los valores. Hay valores que un sujeto no quiere conseguir ni se plantea como importantes para que sus hijos tengan o posean en la vida.

El planteamiento de elección forzosa altera la realidad, esta fue la primera impresión que sacamos de las entrevistas realizadas.

Con las escalas de libertad de elección había otra cuestión que resolver. Los sujetos podían escoger una frase o no escogerla, o por el contrario podían escoger una, dos o ninguna. En las entrevistas vimos de nuevo que lo que más se aproximaba a lo que el sujeto vivía y sentía era una libertad de elección total, de bien escoger una frase del par si así lo veía, o dos si eran significativas para él por igual o no escoger ninguna de las dos si ese era el caso.

Vimos que esta última posibilidad se acercaba más a la realidad del valor y que habría que insistir especialmente en que el sujeto siempre elegía **comparando las dos frases entre sí**, a partir de ahí tenía una decisión de elección total.

3. FIABILIDAD Y VALIDEZ DE LA ESCALA

Una primera validación de la escala ya se llevó a cabo en la elaboración de las frases con el acuerdo interjueces para elegir las frases que corresponderían a cada categoría de valor.

Para comprobar la fiabilidad o consistencia de la medición realizamos una correlación de Pearson entre las dos frases diferentes que hemos utilizado para medir la misma categoría de valor. Si existe una consistencia en la medición cada par de frases de la misma categoría debe correlacionar más alta entre sí que con las frases restantes de las otras categorías.

Las correlaciones entre cada par de frases son todas significativas al 0,01 %.

A continuación exponemos la correlación obtenida para cada par de frases de cada una de las categorías (Los números hacen relación a la frase 1 ó 2 de cada una de las categorías).

Vitales1 con vitales2:	0,3072
Éticos1 con éticos2:	0,5909
Afectivos1 con afectivos2:	0,5207
Trascendentes1 con trascendentes2:	0,7516
Estéticos1 con estéticos2:	0,7073
Sociales1 con sociales2:	0,77
Productivos1 con productivos2:	0,5766
Noéticos1 con noéticos2:	0,4766

Hemos encontrado que se dan correlaciones significativas entre frases de distintas categorías. Esto no debe extrañar ya que los sujetos escogen unos valores por encima de otros, pero no quiere decir que rechacen completamente unos y escojan únicamente los otros, esto es extraño dentro del mundo del valor. Los seres humanos tienden a reconocer a los valores como tales y aunque establecidos dentro de una jerarquía pueden llegar a admitir un abanico muy amplio, sobre todo en el caso de las creencias esto se da de manera más fuerte.

4. RESULTADOS Y CONCLUSIONES

A nivel del acuerdo interjueces de la inclusión de las frases dentro de las categorías se puede considerar que es un elemento de validez de la escala construida.

La fiabilidad de la escala, medida por la correlación entre cada par de frases pertenecientes a la misma categoría, también puede considerarse adecuada ya que todas las frases correlacionan entre sí a un nivel de significación del 1%. Podemos decir, por tanto, que existe una consistencia de la medida de las creencias de valor de los padres con esta escala.

Hemos comprobado que la matriz de correlaciones nos puede servir para entender estructuras de creencias de valor en los padres y cómo estas estructuras están organizadas en el orden social. Recordemos que las creencias de valor están muy influenciadas por grupos de referencia y por el orden social. Lo establecido como «bueno» o «adecuado» por un grupo determinado es lo que se deja entrever en ellas.

Entendemos como estructura de creencias de valor una serie de categorías con una fuerte correlación entre ellas y que marcarían unas guías de conducta determinadas.

El comentario de las correlaciones de estas categorías y las correlaciones que se dan entre algunas frases de distintas categorías no es posible, por su amplitud, incluirlo en este trabajo (consúltese Donoso, 1992). Expondremos únicamente las estructuras de creencias de valor, fuertes relaciones entre categorías de valor, que hemos encontrado como muy significativas:

1. Valores afectivos y vitales.
2. Valores éticos, trascendentes y sociales.
3. Valores trascendentes y éticos.
4. Valores productivos y noéticos.
5. Valores noéticos y éticos

5. BIBLIOGRAFÍA

- BARTOLOMÉ, M. (1981): *Estudio crítico del método de análisis de contenido aplicado a la investigación educativa*. Texto fotocopiado. Universidad de Barcelona.
- DONOSO, T. (1992): *Análisis de los valores en niños de 8 a 10 años*. Tesis doctoral. Universidad de Barcelona.
- GARCÍA HOZ, V. (1972): *Técnicas de trabajo cooperativo en la enseñanza universitaria*. Madrid: ICE. Universidad Complutense.
- GUILIGAN, J. (1976): Beyond morality: Psychoanalytic reflections on shaw, guilt. En Lickona, T. (ed.): *Moral development and behavior theory, research and social issues*. New York: Holt rinehart and winston, pp. 144-159.
- HERS, R., MILLER, J. and FIELDING, G. (1984): *Moral education: A handbook for teachers*. Minneapolis: Winston Press.
- KAY, W. (1976): *El desarrollo moral*. Buenos Aires: Ateneo.

- KURTINES, W. H. and GELWIRTH, J. L. (1984): *Morality, moral and moral development*. Newm York: John Wiley.
- LOCKWOOD, A. (1978): The effects of values clarification and moral development curricula on school age subjects. A critical review of recent research. *Review of educational research*, 48, (3) 325-364.
- MAIER, H. (1982): *tres teorías sobre el desarrollo del niño*: Erikson, Piaget y Sears. Buenos Aires: Amorrortu.
- MALINOWSKY, C. and SMITH, C. P. (1985): Moral reasoning and moral conduct: An investigation prompted by Kohlberg theory. *Journal of personality and social psychology*, 49, (4), 1.016-1.027.
- MASLOW, A. (1962): *Toward a psychology of being*. Princeton: D. Van Nostrand.
- MAY, R. (1978): *El dilema del hombre moderno*. Buenos Aires: Paidós.
- SCHARF, P. (1978): *Readings in moral education*. Minneapolis: Winston Press.

UN INSTRUMENTO PARA EVALUAR LA CONDUCTA EXPLORATORIA EN EL DESARROLLO DE LA CARRERA PROFESIONAL

T. Donoso*, P. Figuera* y M. L. Rodríguez*
Universidad de Barcelona

EL INTERÉS POR EL ESTUDIO DE LA CONDUCTA EXPLORATORIA

Aunque la psicometría y el diagnóstico recibieron más atención que los procesos exploratorios en la historia de la orientación, se detecta, a partir de los años sesenta, una cierta preocupación por definir en qué consiste la exploración de uno mismo y del mundo laboral y cómo va a afectar esta conducta en la elección profesional. Representantes de la psicología experimental (Berlyne, 1960, y Myers, 1941) definieron la conducta exploratoria como el proceso que afecta a la naturaleza de la estimulación que llega a los órganos de los sentidos, con la función de obtener información adicional de ciertos aspectos del entorno. El principal constructo de la teoría de Berlyne es la activación, la actividad, el deseo de realizar una actividad heurística. Esta actuación del orientado arrastra un abanico de conductas de interés para la orientación (inspección, manipulación, observación, etc.) que junto a la disminución de la incertidumbre y de los conflictos, favorecerán la exploración profesional. El mismo interés surgió entre los psicólogos evolutivos quienes introdujeron, entre los estadios vitales, el exploratorio, coincidente, *grosso modo* con la adolescencia. Hoy, existe ya una base para la explotación de la conducta exploratoria a partir de los teóricos del desarrollo de la carrera.

LA CONSTRUCCIÓN DE UNA TEORÍA DE LA CONDUCTA EXPLORATORIA

En las conductas exploratorias han sido contempladas más sus resultados que sus procesos. De ahí que, en los años sesenta, Super y cols. definieran (Jordaan, 1963) el constructo «explorar» como «conducta previa al intento de resolver un problema y a la toma de decisiones.» «(...) que trata de aquellas actividades —mentales o físicas— hechas con un propósito plenamente admitido o inconsciente; con la esperanza de aflorar información acerca de uno mismo del medio; de verificar o conseguir una base para tomar decisiones, concluir, solucionar, construir hipótesis, o para cambiar, sentirse estimulado o responder a los retos.»

De las dimensiones propuestas por Jordaan se han desprendido dos tipos de acciones. Por una parte, la medida de la conducta exploratoria, es decir, el diagnóstico y estimación de la capacidad

* (Dirección de contacto con las autoras: Dep. MIDE, Fac. Pedagogía, C/Baldri i Reixac s/n 3º, D, 08028-Barcelona, T. (93) 3333466, ext. 3384.)

exploratoria; y, por otra, el diseño de programas, más o menos completos, de intervención con el fin primordial de desarrollar en los jóvenes el deseo de explorar. Entre éstas se detectan desde los años sesenta significativos esfuerzos. Por ejemplo, existen paquetes para la exploración profesional que tratan de desarrollar habilidades básicas de comprensión de las ocupaciones y búsqueda de empleos, reales y/o simuladas; uso de los inventarios de preferencias vocacionales y de desarrollo vocacional para enseñar a buscar información y explorar (Aiken y Johnston, 1973); programas para la enseñanza secundaria que pretenden enseñar a los adolescentes y a sus familias a buscar información y a querer completarla con búsquedas complementarias; desarrollo de estilos exploratorios incluyendo la exploración ocupacional en tratamientos orientadores más interales, y así sucesivamente. Las realizaciones más modernas van por el camino de las experiencias profesionales (*work experiences*) y en las intervenciones psicoeducativas muy utilizadas en los programas de educación para la carrera desde una perspectiva evolutiva (Rodríguez Moreno y Figuera, 1992); se ha intentado la creación de materiales autoaplicados facilitadores de la exploración del mundo laboral, aún inéditos y por validar (Rodríguez Moreno, 1989).

Con respecto a los estudios sobre el diagnóstico y estimación de las dimensiones de la conducta exploratoria, se han publicado ya varios instrumentos psicométricos y escalas de estimación que serán el tema del apartado siguiente.

LA ESTIMACIÓN DE LAS DIMENSIONES DE LA CONDUCTA EXPLORATORIA

Tradicionalmente la evaluación específica de la exploración de la carrera ha estado centrada en la dimensión conductual midiendo la frecuencia y la variedad de los actos exploratorios muchas veces con instrumentos reactivos. Sólo recientemente los investigadores han comenzado a valorar la incidencia de los elementos motivacionales e individuales de la exploración (Blustein, 1988a) dirigidos más a la investigación del proceso exploratorio desde una perspectiva cognitiva. Sería el caso de Aiken i Johnston (1973), de Kohen y Breinich (1975), de Sharf (1984) y otros. Entre ellos nos interesa citar las aportaciones de Stumpf, Colarelli y Hartman (1983) y Stumpf y Lockhart (1987) que proponen un modelo integral de la exploración profesional que intenta incluir elementos conductuales y cognitivos propiciatorios de la misma conducta tal y como aparece en la figura 1.

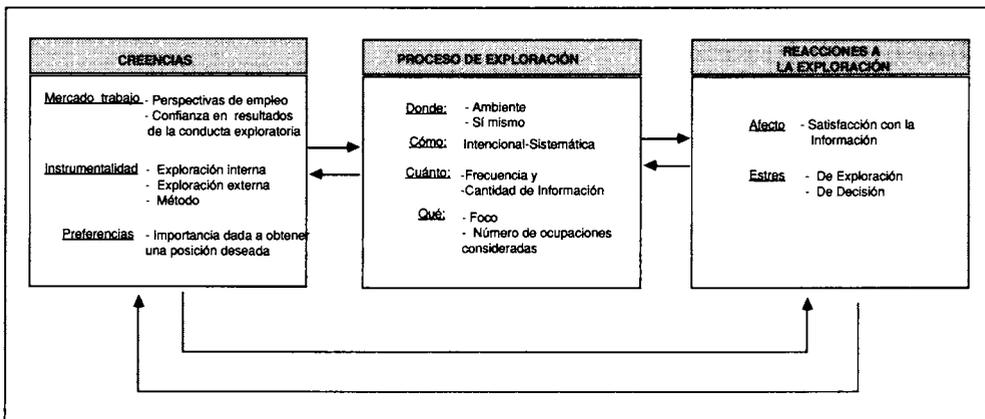


Figura 1. Modelo estructural de exploración profesional. Stumpf, Colarelli y Hartman (1983, p. 194).

Stumpf define la *exploración profesional* como «el conjunto de acciones cognitivas y conductuales que, de una manera intencional, permiten el acceso a la información sobre trabajos, ocupaciones u organizaciones» (1983, p. 258). En el modelo arriba presentado, las conductas son el elemento definitorio del proceso de exploración. Sin embargo, el valor que da el sujeto a la información obtenida —o a aquella que esperaba haber obtenido— determina las reacciones afectivas hacia el propio proceso de exploración. De ese modo, y tal como aparece representado en la figura 1, las creencias o expectativas del actor acerca del valor o la utilidad de su futura conducta exploratoria están determinadas por la experiencia pasada. De ello se deducen tres elementos base del modelo; a saber:

1. El Proceso de Exploración

Este primer elemento del modelo delimita la dimensión conductual del constructo, concretada en una serie de componentes esenciales del proceso: (a) *Dónde* se realiza la exploración (fuente de los datos); (b) *Cómo* se explora: (conducta puede ser sistemática e intencional o fortuita y resultado del azar). (c) *Cuánto se explora*: (la frecuencia de la acción de explorar y la cantidad de información que la persona obtiene como resultado de su conducta). Y (d) *Qué se busca*: (número de áreas u ocupaciones exploradas y a la intensidad exploratoria).

Diferentes investigaciones resaltan la importancia del proceso exploratorio en los logros posteriores de la carrera: De hecho existe mucha literatura prescriptiva que demuestra la necesidad de explorar para ingresar con más garantías en el mundo del trabajo. Stumpf, Austin y Hartman (1984) detectan que, el objeto de la exploración y la cantidad de información obtenida estaban significativamente relacionados con dos índices objetivos de inserción profesional: ejecución en la entrevista de selección ocupacional y número de ofertas obtenidas. Además de facilitar la entrada, la frecuencia e intensidad de la exploración, también favorecen la salida de situaciones de desempleo continuado, y el ajuste laboral posterior (Vid. el barrido que hace Blustein al respecto, 1988b). Estos resultados han sido replicados en nuestro país en el marco de la teoría de las preferencias profesionales (Muñoz Sastre, 1992) y de la psicología del desempleo (Blanch, 1990, Álvaro 1992). Parece ser, además, que los componentes conductuales de la exploración profesional están asociados a indicadores de, demostraron que aquellas personas más implicadas en el proceso de exploración de carrera mostraban unos niveles más altos de satisfacción laboral.

2. Reacciones a la Exploración

Definida esta dimensión como sentimientos o creencias acerca de la información *obtenida* o *no obtenida*, éstos van a tomar diferentes formas que mediatizarán, a su vez, las expectativas y la conducta de búsqueda posterior. Entre las reacciones que el orientado puede experimentar como consecuencia de los resultados percibidos de su conducta se detecta que: (a) El *grado de satisfacción* con los resultados de la exploración puede variar tanto la dirección y el foco de la búsqueda, como la sistematicidad de la misma; (b) *La ansiedad o estrés* vividos en el proceso, son el resultado de apreciar algún grado de incertidumbre respecto a las posibilidades de logro de los objetivos deseados. Tanto las situaciones de insatisfacción como de estrés influyen en la conducta posterior: desencadenan en la persona una serie de mecanismos de afrontamiento (entre los que puede incluirse la misma evitación del resultado) que podrían implicar eliminar la conducta de exploración (Vid. la literatura del desempleo, Albaracín, 1991, Blanch, 1990 y Pérez Escoda, 1992, entre otros muchos).

3. Creencias acerca de la conducta exploratoria

Aceptar el principio de que la exploración es un proceso cognitivo/consciente implica que, en un momento dado, los individuos están actuando bajo un conjunto de creencias y percepciones; y que éstas independientemente de que sean reales o no, determinarán su conducta. En la exploración profesional tres áreas de creencias adquieren relevancia especial: (a) La percepción sobre las *condiciones del mercado de trabajo* (específicamente sobre las perspectivas de empleo y, más concretamente la percepción de la posibilidad de conseguir las metas profesionales. (b) La creencia en la *instrumentalidad o utilidad* de las conductas y métodos de búsqueda o exploración utilizados influirán sobre el dónde y el cómo busca la persona, porque se constituyen en elementos de motivación (Vid. las teorías de Vroom, 1964, Wanous, 1980, y los trabajos de Deci y Ryan, 1980). Las personas que creen en la utilidad de una determinada conducta estarán más dispuestas a invertir esfuerzo que si piensan que es una acción inútil. Y, (c) La *importancia* dada por la persona a las *metas profesionales* o a un determinado aspecto en concreto. Por un lado cabe esperar que en aquellas personas más interesadas en conseguir sus metas, el nivel de implicación en la exploración será mayor, ya que los patrones de exploración pueden diferir de acuerdo a la importancia que la persona le concede más a un objeto de exploración profesional que a otro.

ADAPTACIÓN DE LA ESCALA CES DE STUMPF Y COLABORADORES

Las especificaciones de las dimensiones de la exploración de la carrera del modelo de Stumpf y otros (1983) se desarrollan en el siguiente listado. La numeración de cada especificación remite a la lectura del cuadro 2.

Proceso de exploración

1. **EXPLORACIÓN AMBIENTAL:** Extensión de la exploración de la profesión examinando ocupaciones, trabajos y organizaciones en los últimos tres meses.
2. **AUTO-EXPLORACIÓN:** Extensión de la exploración de la carrera que comprende auto-valoración y retrospección en los últimos tres meses.
3. **NÚMERO DE OCUPACIONES CONSIDERADAS:** Número de diferentes áreas ocupacionales en las que uno está adquiriendo información.
4. **EXPLORACIÓN INTENCIONAL Y SISTEMÁTICA:** Extensión con que uno adquiere información sobre uno mismo y el ambiente de una manera sistemática.
5. **FRECUENCIA:** Promedio de veces por semana que uno busca información profesional en un periodo de dos meses.
6. **CANTIDAD DE INFORMACIÓN:** Cantidad de información adquirida sobre ocupaciones, trabajos, organizaciones y uno mismo.
7. **FOCO:** Seguridad de la preferencia por una ocupación particular, trabajo y organización.

Reacciones a la exploración

8. **SATISFACCIÓN CON LA INFORMACIÓN:** Satisfacción con la información obtenida respecto a ocupaciones, trabajos y organizaciones relativas a los propios intereses, capacidades y necesidades.
9. **TENSIÓN EXPLORATORIA:** Cantidad de tensión no deseable en relación a otros eventos significativos de la vida.

10. TENSION DECISORIA: Cantidad de tensión no deseable, en relación a otros eventos significativos de la vida,

Expectativas

11. PERSPECTIVA PROFESIONAL: Cuán favorables son las posibilidades que uno busca en un área profesional.
12. CERTEZA DE LOS RESULTADOS DE LA EXPLORACIÓN PROFESIONAL: Grado de certeza sentido de obtener un lugar de trabajo deseado.
13. INSTRUMENTALIDAD DE LA BÚSQUEDA EXTERNA: La probabilidad que explorar el ambiente sobre oportunidades profesionales conducirá a obtener metas profesionales.
14. INSTRUMENTALIDAD DE LA BÚSQUEDA INTERNA: Probabilidad de que reflexionar retrospectivamente sobre conductas profesionales pasadas conducirá a conseguir las metas profesionales.
15. INSTRUMENTALIDAD DEL MÉTODO: Probabilidad de que la exploración profesional intencional y sistemática conducirá a conseguir las metas profesionales.
16. IMPORTANCIA DE OBTENER UN PUESTO DE TRABAJO PREFERIDO: Grado de importancia que se da al hecho de obtener una profesión preferente.

En el cuadro adjunto se observan las diferencias entre la escala original y la escala adaptada en cuanto al número de ítems y al lugar que ocupa el ítem en la escala.

CUADRO 2

ESPECIFICACIONES DE LAS DIMENSIONES E ÍTEMS DE LA ESCALA ORIGINAL Y ADAPTADA

ITEMS ESCALA GENERAL	Proceso de exploración							Reacciones a la exploración			Expectativas					
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Nº de ítems	6	5	1	3	1	3	5	6	3	4	3	3	3	4	4	5
Lugar del ítem en la escala	18-23	24-28	29	15-17	36	1-3	10-14	4-9	53-55	56-59	30-32	33-35	45-47	41-44	37-40	48-52
ITEMS ESCALA ADAPTADA																
Nº de ítems	6	5	1	3	1	3	3	4	3	4	3		3	3	4	4
Lugar del ítem en la escala	13-18	20-24	25	11-13	29	1-3	8-10	4-7	44-46	47-50	26-28		37-39	34-36	30-33	40-43

Cuadro 2 : Especificaciones de las dimensiones e ítems de la escala original y adaptada.

La escala adaptada reduce el número de ítems original, de 59 a 50. Los criterios para realizar las reducciones han sido, en la mayoría de casos, terminológicos y contextuales. Así la dimensión **foco** ha pasado de cinco ítems a tres. Igual ocurre con **satisfacción con la información**, que pasa seis ítems a cuatro. **Instrumentalidad de la búsqueda interna** pasa de cuatro a tres ítems e **importancia de obtener un puesto de trabajo preferido** de cinco a cuatro ítems. La dimensión **certeza de los resultados de la exploración profesional** queda subsumida en la de **perspectiva profesional**, ya que

la primera era demasiado repetitiva y podía producir desorientación y cansancio. Así la escala de 50 ítems que ha resultado agiliza la lectura y se adapta mejor a nuestro contexto. Nuestros planes, a un plazo medio, continúan en la línea de la fiabilización y validación de esta escala de Conducta Exploratoria (ESCE).

REFERENCIAS

- AIKEN, J., JOHNSTON, J. A. (1973): Promoting Career Information Seeking Behaviors in College Students, *Journal of Vocational Behavior*, 3, 1, 81-87.
- ÁLVARO, J. L. (1992): Desempleo y bienestar psicológico. Madrid: Siglo XXI.
- ALBARRACÍN, J. (1991): Composición y duración del desempleo, S. BENTONILLA y L. TOHARRIA (Comp.): *Estudios de Economía del Trabajo en España: El problema del paro!* Madrid. Ministerio de Trabajo y Seguridad Social. 143-170.
- BERLYNE, D. E. (1960): *Conflict, arousal and curiosity*, New York: McGraw-Hill Book Company.
- BLANCH, J. M. (1990): *Del viejo al nuevo paro. Un análisis psicológico y social*. Barcelona: PPU.
- BLUSTEIN, D. L. (1988a): The relationship between motivational processes and career exploration, *Journal of Vocational behavior*, 33, 340-357.
- BLUSTEIN, D. L. (1988b): Individual and contextual factors in career exploration, *Journal of Vocational Behavior*, 33, 203-216.
- DECI, E. L. y RYAN, R. M. (1985): *Intrinsic motivation and self-determination in human behavior*, New York: Plenum.
- JORDAAN, J. P. (1963): Exploratory Behavior: The formation of self and occupational concepts, en D. E. Super, *Career development: Self-concept theory*, New York: College Entrance Examination Board.
- KOHEN, A. J. y BREINICH, S. (1975): Knowledge of the world of work: A test of occupational information for young men. *Journal of Vocational behavior*, 6, 133-144,
- MUÑOZ SASTRE, M. T. (1992): Calidad de la Información profesional y 'disponibilidad' de índices sociales, *Revista de Psicología General y Aplicada*, 45, 2, 161-167.
- MYERS, G. E. (1941): Principles and techniques of vocational guidance, New York: McGraw-Hill Bo. Co. Inc.
- PÉREZ ESCODA, N. (1992): *Parados de larga duración: Análisis de perfiles*. Comunicación presentada al V Seminario Iberoamericano de Orientación Escolar y Profesional. AEOEP. Tenerife, 13-16 de mayo.
- RODRÍGUEZ MORENO, M. L. (1989-91): *Guía para explorar el mundo laboral*, Compuesta de diez unidades didácticas y tests de instrucción, por objetivos. Investigación financiada por la UNESCO y la CIRIT, Inédita. Universidad de Barcelona.
- RODRÍGUEZ, M. L., FIGUERA, P. (1992-93): *La educación para la carrera profesional*, Programa de la nueva materia en el currículum de formación del orientador. Facultad de Pedagogía: Departamento MIDE, Universidad de Barcelona.
- SHARF, R. S. (1984): Vocational information seeking behavior: another view, *The Vocational Guidance Quarterly*, 33, 2, 120-129.
- STUMPF, S. A., COLARELLI, S. M. y HARTMAN, K. (1983): Development of the Career Exploratory Survey (CES), *Journal of Vocational Behavior*, 22, 26, 191-227.
- STUMPF, S. A., BRIEF, A. P. y HARTMAN, K. (1987): Self-efficacy Expectations and coping with Career-related events, *Journal of Vocational Behavior*, 31, 91-108.
- STUMPF, S. A. y LOCKHART, M. C. (1987): Career exploration: Work-role salience, work preferences, beliefs, and behavior, *Journal of Vocational Behavior*, 30, 258-269.

- SUPER, D. E. y HALL, D. T. (1978): Career development: exploration and planning, *Annual Review of Psychology*, 29, 333-372.
- WANOUS, J. P. (1980): *Organizational entry: recruitment, selection and socialization of new comers*, Reading, Mass.: Addison-Wesley.
- WROOM, V. H. (1964): *Work and Motivation*, New York: Wiley.

ALTERNATIVA AL «CHILDREN'S DEPRESSION INVENTORY» (CDI): PROCEDIMIENTO Y METODOLOGÍA

L. Cabrera Pérez, y P. García-Medina
La Laguna/Tenerife

RESUMEN

Nuestra propuesta, consiste en dotar al Cuestionario de Depresión Infantil (Children's Depression Inventory, CDI) de una mayor validez de contenidos, puesta en cuestión por algunos autores. Para ello se trata de extraer los contenidos Cuestionario de Depresión de Beck, para adultos (Beck Depression Inventory, BDI); y readaptar los contenidos del CDI. Nos sirvió una muestra de 220 niños y adolescentes. El análisis correlacional inter-ítems e inter-instrumentos nos proporcionó los contenidos nucleares de un nuevo test (TED). A sus contenidos se añadieron otros considerados en la literatura psicopatológica y no contemplados en ninguno de los dos instrumentos utilizados para la elaboración de éste.

JUSTIFICACIÓN DE LA INVESTIGACIÓN

El desarrollo de técnicas psicométricas para detectar y medir la sintomatología depresiva en niños y adolescentes, es un fenómeno que comenzó con cierta preocupación en los años 70 (p.e., Kovacs y Beck) pero que se ha mostrado poco sensible a la incorporación de nuevas alternativas que optimicen la medición científica de la psicopatología depresiva.

El «*Children's Depression Inventory*» (Kovacs, U.S.A., 1977) es el primer instrumento psicométrico que se elabora siguiendo un criterio de homogenización de sus ítems y cuya finalidad es la aplicación a niños y adolescentes. Este es una adaptación, hecha a partir del cuestionario de depresión para adultos de Beck («*Beck Depression Inventory*», BDI, 1967). Su justificación radica en la presumible sintomatología común para la niñez, la adolescencia y la adultez.

El CDI consta de 27 ítems, cada uno de los cuales con tres respuestas alternativas en orden de frecuencias. Recoge una amplia gama de síntomas depresivos como son: a) las alteraciones del humor; b) la capacidad hedónica; c) las funciones vegetativas; c) la autoevaluación; y d) otras conductas de relaciones interpersonales (Polaino-Lorente, 1988). Se toma como una medida aceptable de *screening*, y con poder discriminativo cuando se aplica a población «normal» escolarizada. Pero, desde nuestro punto de vista, su mayor utilidad reside en su capacidad para medir la severidad del síndrome depresivo (Kovacs, 1983). Si bien han sido varios los trabajos posteriores que se han realizado sobre sus propiedades psicométricas, y su uso es bastante generalizado, consideramos que aún está en período de prueba.

Los estudios sobre su fiabilidad y validez han sido numerosos (Friedman-Butler, 1979; Green, 1980; Kovacs, 1983; Seligman y cols., 1984; Mullins y cols., 1985; Lovobits y cols., 1985; Mestre y cols., 1992). Los resultados, sin embargo, han sido moderadamente consistentes en los distintos trabajos. Lo cual da idea de que el instrumento puede mejorar algunos de sus aspectos. Los autores atribuyen esta falta de consistencia a características particulares de los diferentes estudios. Una de las dificultades más serias está en la reiterada aplicación a muestras muy reducidas, lo cual limita su generalización. Y por otro lado tenemos otra dificultad añadida. La aplicación centrada casi siempre, salvo contados casos (p.e., Kovacs, 1983) a población sólo clínica o sólo normal, lo cual limita su aplicabilidad como instrumento epidemiológico. Otros autores (Gross y Wixted, 1988) sin embargo, sugieren una menor consistencia para el CDI. Las limitaciones parecen derivarse de: a) su limitada validez de contenido (p.e., «he fracasado totalmente», o «me siento mala persona y despreciable» no son contenidos que se incluyen); b) su relativa restricción en la intensidad de la sintomatología (p.e. los niños pueden perfectamente discriminar entre «nunca», «de vez en cuando», «mucho» y «siempre»); c) no da opción a recoger en ningún caso el tiempo que ha durado el estado emocional; y d) Bellack and Hersen (1988) sugieren además que los instrumentos como el CDI, debieran poder usarse como recurso para identificar no sólo depresión sino también otros problemas emocionales. Superando estas dificultades, sí que servirían como medidas válidas de *screening*.

En cuanto a la consistencia, los datos obtenidos muestran mayor solidez. El valor promedio para diferentes consistencias internas halladas para el CDI muestran un Alfa de Cronbach de 0,84 (Síntesis hecha en base a datos recopilados por Polaino-Lorente, 1988).

SUPERANDO LAS DIFICULTADES

Ante las dificultades anotadas anteriormente, nos proponemos dar salida a este *impase* dotando a este instrumento de otras características.

En primer lugar hemos partido del relativo acuerdo entre investigadores y clínicos de que hay rasgos comunes entre la sintomatología depresiva de los adultos y de los niños. Por ellos procedimos a extraer aquellos contenidos convergentes para el BDI y el CDI. De este modo pudimos observar que algunos elementos faltaban en el segundo de los instrumentos. La forma de proceder la dividimos en tres fases: 1) aplicación a una muestra de adolescentes del BDI y del CDI (n=100; media de edad 15); 2) elaboración de un nuevo cuestionario, el TED (Test Experimental de Depresión); y 3) nueva aplicación a una muestra de niños del TED junto con el CDI (n=120; edad media 10 años).

CONSTRUCCIÓN DEL TED COMO ALTERNATIVA

Partiendo de los datos obtenidos en el primer pase del CDI y el BDI, se hizo un análisis de correlaciones. Se situó el punto de corte en valores «r» iguales o mayores que 0,40. A partir de ahí se reelaboraron nuevos ítems. Tales ítems tenían las siguientes características: 1) recogían 4 intensidades (en lugar de los 3 que recoge el CDI) ; 2) incluían contenidos procedentes de ambos cuestionarios; 3) se respetaron casi todos los contenidos del CDI; 4) se incluyeron nuevos aspectos no contemplados ni el BECK ni en el CDI (p. ej., la intensidad de tiempo de duración de la sintomatología depresiva).

Una vez construido el instrumento se puso a prueba en una muestra de población infantil junto al CDI. Los resultados hallados fueron los siguientes: 1) la consistencia interna de ambos instrumentos fue un Alfa de Cronbach para el CDI=0,90, y para el TED=0,72 (n=120, de niños de 10 años de edad media).

TABLA N.º 1
 ÍTEMS DE LOS CUESTIONARIOS BDI Y CDI INTERRELACIONADOS

<p>CDI-1. Estoy triste de vez en cuando. Estoy triste muchas veces. Estoy triste siempre.</p>	<p>Beck-5. $r = .46$ (tengo siempre como una pena encima que no me la puedo quitar). Beck-47. $r = .54$ (me siento irritado todo el tiempo). Beck-49. $r = .54$ (He perdido todo mi interés por los demás y no me importan en absoluto). Beck-2. $r = -.48$ (no me encuentro triste). Beck-40. $r = -.41$ (no tengo pensamientos de hacerme daño). Beck-41. $r = -.41$ (no lloro más de lo habitual).</p>
<p>CDI-5. Soy malo/a siempre. Soy malo/a muchas veces. Soy malo/a algunas veces.</p>	<p>Beck-11. $r = -.66$ (he fracasado totalmente como persona).</p>
<p>CDI-9. No pienso en matarme. Pienso en matarme pero no lo haría. Quiero matarme.</p>	<p>Beck-35. $r = .43$ (tengo pensamientos de hacerme daño, pero no llegaría a hacerlo). Beck-36. $r = .41$ (siento que estaría mejor muerto). Beck-69. $r = .47$ (me canso más fácilmente que antes). Beck-40. $r = -.69$ (no tengo pensamientos de hacerme daño). Beck-71. $r = -.45$ (no me canso más de lo habitual).</p>
<p>CDI-10. Tengo ganas de llorar todos los días. Tengo ganas de llorar muchos días. Tengo ganas de llorar de vez en cuando.</p>	<p>Beck-23. $r = -.51$ (me siento prácticamente todo el tiempo mala persona y despreciable).</p>
<p>CDI-16. Todas las noches me cuesta dormir. Muchas noches me cuesta dormir. Duermo muy bien.</p>	<p>Beck-2. $r = .41$ (no me encuentro triste).</p>
<p>CDI-18. La mayoría de los días no tengo ganas de comer. Muchos días no tengo ganas de comer. Como muy bien.</p>	<p>Beck-76. $r = .43$ (tengo el mismo apetito de siempre). Beck-74. $r = -.40$ (Mi apetito no es tan bueno como antes).</p>
<p>CDI-19. No me preocupa el dolor ni la enfermedad. Muchas veces me preocupa el dolor y la enfermedad. Siempre me preocupa el dolor y la enfermedad.</p>	<p>Beck-83. $r = -.43$ (no me preocupa mi salud más de lo normal).</p>
<p>CDI-20. Nunca me siento solo/a. Me siento solo/a muchas veces. Me siento solo/a siempre.</p>	<p>Beck-2. $r = -.40$ (no me encuentro triste). Beck-40. $r = -.43$ (no tengo pensamientos de hacerme daño).</p>
<p>CDI-25. Nadie me quiere. No estoy seguro/a de que alguien me quiera. Estoy seguro de que alguien me quiere.</p>	<p>Beck-5. $r = -.51$ (tengo siempre como una pena encima que no me la puedo quitar). Beck-11. $r = -.48$ (he fracasado totalmente como persona). Beck-23. $r = -.54$ (me siento prácticamente todo el tiempo mala persona y despreciable). Beck-38. $r = -.48$ (tengo planes decididos de suicidarme). Beck-47. $r = -.50$ (me siento irritado (enfadado) todo el tiempo).</p>
<p>CDI-27. Me llevo bien con la gente. Me peleo muchas veces. Me peleo siempre.</p>	<p>Beck-11. $r = .52$ (he fracasado totalmente como persona —como padre, marido, hijo/a, estudiante, madre, etc.).</p>

CONCLUSIONES

Respecto al CDI encontramos como el TED se resiente en cuanto a la consistencia interna. Si bien en principio se muestra como limitación, ésta se ve compensada si tenemos en cuenta que una de las causas es la mayor heterogeneidad de los contenidos del TED. Sin embargo, muestra como ventaja respecto al CDI sus mejoras en cuanto a diseño. Así los niveles de intensidad de los ítems añadidos, y la contabilización del tiempo de duración del síntoma proporcionan nuevos recursos para la selección y corte (screening).

BIBLIOGRAFÍA

- BECK, A. T. (1969): *Depression Clinical, experimental and theoretical aspects*. New York: Hoeber.
- BELLACK, A. S. y HERSEN, M. (1988): *Behavioral Assessment: A Practical Handbook*. Nueva York: Pergamon Press.
- DEL BARRIO, V. (1990): Situación actual de la evaluación de la depresión infantil. *Evaluación Psicológica*, 2 (Vol. 6), 171-209.
- FRIEDMAN, J. J., BUTLER, L. F. (1979): *Development and evaluation of a test battery to assess childhood depression*. Final report to Health and Welfare. Canadá.
- GREEN, B. J. (1980): Depression in early adolescence: and exploratory investigation of its frequency, intensity and correlates. *Dissertation Abstracts International*, 41, 3890-B.
- GROSS, A. M., y WIXTED, J. T. (1988): Assessment of child behavior problems. En A. S. BELLACK y M. HERSEN (Eds.): *Behavioral Assessment: A Practical Handbook*. Nueva York: Pergamon Press.
- GROSS, A. M., y WIXTED, J. T. (1988): Assessment of child behavior problems. En A. S. BELLACK y M. HERSEN (Eds.): *Behavioral Assessment: A Practical Handbook*. Nueva York: Pergamon Press.
- KOVACS, M. (1983): *The Children's Depression Inventory: a self-rated depression scale for school-aged youngster*. University of Pittsburgh School of Medicine. Documento no publicado.
- KOVACS, M., BECK, A. T. (1977): An empirical-clinical approach toward a definition of childhood depression. En: SCHULTERBRANDT, J. G., RASKIN, A. (Eds.): *Depression in childhood: diagnosis, treatment and conceptual models*. Raven.
- LOVOBITS, D. A., HANDAL, P. J. (1985): Childhood depression-Prevalence using DSM-III criteria and validity of parent and child depression scales. *Journal of Pediatric Psychology*, 10.
- MESTRE, V., FRÍAS, D., GARCÍA-ROS, R. (1992): Propiedades psicométricas del Children's Depression Inventory (CDI) en población adolescente: Fiabilidad y Validez. *Psicológica*, 13, 149-159.
- MULLINS, L. L., SIEGEL, L. J., HODGES, K. (1985): Cognitive problems-solving and life events correlates of depressive symptoms. *Journal of Abnormal Child Psychology*, 13.
- POLAINO-LORENTE, A. (1988): *Las Depresiones Infantiles*. Madrid: Morata.
- SELIGMAN, M. E. P., PETERSON, C., KASLOW, N. J., TANENBAUM, R. L., ALLOY, L. B., ABRAMSON, L. Y. (1984): Attributional style and depressive symptoms among children. *Journal of Abnormal Psychology*, 93.

COMUNICACIONES A LA PONENCIA IV

LA TEORÍA DE LA INFORMACIÓN EN EL ANÁLISIS DE DATOS CUALITATIVOS

Xabier de Salvador González
Universidade da Coruña

A partir de los trabajos de Shannon-Weaver y McMillan a mediados de este siglo, se desarrolló la llamada **Teoría matemática de la Información** con origen en el concepto termodinámico de *desorden* o *entropía*. Los conceptos y resultados de esta teoría han tenido un amplio dominio de aplicaciones en el campo de las ciencias humanas y sociales (lingüística, psicología,...). Si bien ha sido R. A. Fisher quien introduce en 1925 el término de *información* en estimación estadística (cantidad de información proporcionada por los datos sobre un parámetro no conocido), las ideas de Shannon-Weaver, más comprometidas con la probabilidad, también fueron asumidas en estadística.

Nos interesa aquí esta teoría en su aplicación al análisis de datos cualitativos. Fueron varios los trabajos realizados bajo este objetivo. Entre ellos destacan los de McGill (1954) y Kullback (1968) en análisis multivariable de tablas de contingencia para medidas y tests de asociación, Garner (1956) en análisis de varianza, Krippendorff (1982,1986) en análisis de regresión y modelos estructurales.

Todos estos estudios parten de los conceptos básicos de entropía y cantidad de información que una variable transmite a otra definidos por Shannon-Weaver:

Para una variable discreta X con m categorías, se define la **entropía** de X, representada por $H(X)$, como medida de incertidumbre, por la expresión

$$H(X) = -\sum_{i=1}^m (P_i \cdot \log_2 P_i)$$

siendo P_i la probabilidad de que la variable X tome el valor i [$P[X=i]$], con $i=1, \dots, m$. Esta medida del grado de indeterminación que presenta una variable es no-negativa y presenta su valor máximo en una distribución equiprobable de sus valores (distribución uniforme), aumentando con el número de éstos. Se recoge así el *principio de razón insuficiente* (o principio de indiferencia) utilizado ya en el siglo XVII por Bernouilli y establecido por Laplace en el siglo XIX: supuesto que no hay fundamento para preferir una modalidad a otra de un acontecimiento, debemos considerar que tienen la misma probabilidad de ocurrencia.

Considerando dos variables discretas X (con categorías i) e Y (con categorías k) y su tabla de contingencia X*Y, podemos definir, aparte de sus respectivas entropías (entropías marginales), las entropías condicionales (bajo las distribuciones de X condicionada a Y, e Y condicionada a X) y la entropía total (bajo la distribución conjunta de ambas variables):

entropía condicionada:

$$H(Y|X) = \sum_{i=1}^m \sum_{k=1}^n (P_{ik} \cdot \log_2 P_{k|i})$$

entropía total:

$$H(XY) = \sum_{i=1}^m \sum_{k=1}^n (P_{ik} \cdot \log_2 P_{ik})$$

siendo P_{ik} la probabilidad conjunta $P[X=i, Y=k]$ y $P_{k|i}$ la probabilidad condicionada $P[Y|X=i]$, con $P_{ik} = P_{k|i} \cdot P_i$.

La propiedad de la disminución de la entropía

$$H(Y|X) \leq H(Y)$$

pone de manifiesto para Gil Álvarez (1981, p. 29) que «el conocimiento de una experiencia sólo puede disminuir nuestra incertidumbre, nunca aumentarla. Además, la reducción será efectiva salvo en el caso en que ambas experiencias no tengan ninguna relación».

Es precisamente esta propiedad la que da lugar a Shannon a definir la **cantidad de información** que una variable transmite a otra, representada por $I(XY)$, como la diferencia entre la entropía marginal y su entropía condicionada:

$$I(XY) = H(Y) - H(Y|X)$$

es decir, la disminución de la incertidumbre de una variable que se consigue con el conocimiento de otra. Esta cantidad se mide en «bits» [$1 \text{ bit} = -\log_2(1/2)$] que para McGill (1954, p. 99) «representa la información aportada por la elección entre dos alternativas igualmente probables».

Si consideramos la variable X como variable antecedente (transmisor) de Y (receptor), la ecuación anterior da pie a descomponer aditivamente la entropía de Y en dos componentes: la información transmitida por el emisor, $I(XY)$, y la entropía condicionada $H(Y|X)$ que se la denomina *ruido*. La dependencia total entre las variables podemos asociarla a una *transmisión sin ruido*: cada código de entrada (categorías de X) está asociado biunívocamente a un código de salida (categorías de Y). En el otro caso extremo, la independencia entre variables, viene dada por un *canal inútil*, incapaz de transmitir información: cada código de entrada tiene la misma probabilidad de estar asociado a cualquiera de los códigos de salida.

Considerando un sistema de tres variables X , T e Y podemos atender a diferentes tipos de cantidades de información:

- *informaciones bivariadas*: las producidas entre dos variables,
- *informaciones condicionales*: la información transmitida entre dos variables condicionada a una tercera, y definida por la expresión

$$I(XY|T) = H(Y|T) - H(Y|XT)$$

— *información total*: la cantidad de información que se transmite en el sistema y definida como la entropía máxima de éste (cuando se produce la independencia entre todas las variables) y su entropía observada (la de la distribución conjunta de todas las variables)

$$I(XTY) = H(X) + H(T) + H(Y) - H(XTY)$$

Esta información total se puede descomponer en informaciones binarias:

$$I(XTY) = I(XT) + I(TY) + I(XYT)$$

Las diferentes cantidades de información cuando se calculan sobre una muestra representativa de la población pueden ser probadas en significación a partir de la razón de verosimilitud (Miller and Madow, 1954):

$$L^2 = 1.3863 \cdot N \cdot I(XY)$$

siendo N el tamaño muestral. Para muestras grandes, L^2 se distribuye aproximadamente como χ^2 con $(i-1) \cdot (k-1)$ g.l.

Como señalamos al principio, estas ideas fueron utilizadas para realizar diferentes tipos de análisis multivariantes con datos cualitativos (variables nominales). Así para un modelo tridimensional de transmisión de información con variable consecuente Y y antecedentes X, T, se puede descomponer la entropía de Y (McGill, 1954) en:

$$H(Y) = H(Y|XT) + I(XY) + I(TY) + A(XTY)$$

donde la entropía condicionada de Y a sus variables antecedentes X y T, $H(Y|XT)$, supondría la variabilidad residual o no-explicada en Y (término de error) y $A(XTY)$ el término de interacción entre las tres variables que representaría la ganancia (o pérdida) en la transmisión de la información en una de las informaciones bivariadas, $I(XY)$ o $I(TY)$, debido al conocimiento adicional de la tercera variable (T o X respectivamente).

Si bien los análisis de la varianza en variables nominales con la utilización de la teoría de la información fueron desplazados por los modelos log-lineales, tanto el análisis de regresión como el análisis causal realizados con dicha teoría presentan ciertas ventajas sobre los modelos logit:

— permite trabajar con politomías en la variable dependiente frente a la necesaria dicotomía en regresión logística,

— asigna un único valor a los caminos causales cuando se trabajan con variables politómicas frente a los múltiples valores que asignaría el análisis causal a partir de modelos logit.

Hoy en día, el uso de las diversas técnicas de análisis para este tipo de datos (análisis de correspondencias, modelos log-lineales, análisis multivariantes no-lineales,...) suponen un importante reto en el campo de la investigación educativa.

BIBLIOGRAFÍA

- ASH, R. (1965): *Information theory*. New York: John Wiley & Sons.
- GARNER, W. R. (1956): «The relation between information and variance analyses». *Psychometrika*, 21 (3); 219-228.
- GIL ÁLVAREZ, P. (1981): *Teoría matemática de la información*. Madrid: Edic. ICE.
- GUIASU, S. et THEODORESCU, R. (1966): *La Théorie mathématique de l'information*. París: Dunod, 1968.
- KRIPPENDORFF, K. (1982): «Regression analysis using information theory». En L. TRONCALE (Ed.): *A Survey of Systems Methodology*. Louisville, KY: Society for General System Research; 1.007-1.012.

- KRIPPENDORFF, K. (1986): *Information theory. Structural models for qualitative data*. Beverly Hills, CA: Sage.
- KULLBACK, S. (1968): *Information theory and statistics*. Gloucester, Mass.: Peter Smit, 1978.
- KULLBACK, S. and KEEGEL, J. C. (1984): «Categorical data problems using information theory». En P. R. KRISHNAIAH & P. K. SEN (Ed.): *Handbook of Statistics, 4: Nonparametric methods*. Netherland: North-Holland; 831-871.
- McGILL, W. J. (1954): «Multivariate information transmission». *Psychometrika*, 19 (2), 97-116.
- MILLER, G. A. and MADOW, W. G. (1954): On the maximum likelihood estimate of the Shannon-Weaver measure of information. Report No. TR-54-75. Washintong, DC: Air Force Cambridge Research Center.
- SHANNON, C. E. and WEAVER, W. (1949): *The mathematical theory of communication*. Urbana: University of Illinois Press.

POSIBILIDADES DE WORDPERFECT 5.1. PARA EL ANÁLISIS DE DATOS CUALITATIVOS

Luis Carro San cristóbal
Universidad de Valladolid

INTRODUCCIÓN

La recogida de datos en una investigación cualitativa significa considerar cualquier dato como unidad con significado propio susceptible de ser analizado. A priori, todo dato puede ser significativo y útil. Esto conlleva reunir grandes cantidades de datos que más tarde nos van a suscitar algunos interrogantes: «y ahora, ¿qué hago con todo esto?, ¿para qué me sirve?, ¿cómo voy a analizar estos datos?».

En los últimos años estamos asistiendo a una gran proliferación de trabajos de investigación en los que los datos recogidos directamente del campo de estudio ya no son traducidos en números, sino que las aportaciones de las personas son tratadas como datos más de carácter cualitativo que cuantitativo. Estos datos están presentados en citas textuales de las opiniones, puntos de vista, apreciaciones, documentos personales, etc. que son recogidos a través de entrevistas, notas de campo, descripciones ecológicas de la observación directa en el campo de trabajo.

La existencia de los actuales ordenadores personales ha permitido de forma notable e importante agilizar el trabajo y reducir los esfuerzos de trabajar con los datos de nuestras investigaciones. La aplicación del software informático al proceso de análisis de datos ha supuesto una gran contribución en la medida que ayuda a sistematizar, organizar y reducir gran cantidad de datos, ya sean éstos de carácter numérico (para su posterior tratamiento estadístico) o en formato de texto escrito (caracteres alfanuméricos, en términos informáticos).

El uso común del ordenador para el trabajo personal ha supuesto una gran proliferación de programas informáticos desarrollados con la finalidad de agilizar nuestro trabajo. A diferencia de las ciencias exactas que trabajan con programas de carácter matemático y estadístico, en ciencias sociales es utilizado el procesador de textos como una potente herramienta de trabajo. Esto, unido a un conocimiento de la lógica del procesador de texto habitual, permite tomar decisiones de cuando y cómo seleccionar un programa más especializado (REID, 1992:145).

Para el análisis de datos cualitativos existen diferentes programas informáticos que a veces puede resultar difícil su localización y la adquisición por parte del usuario, como son LispQual (DRASS, 1980), QualPro (BRACKMAN, 1987), KEDIT (MANSFIELD SOFT.GROUP, 1987), The ETHNOGRAPH (SEIDEL y otros, 1988) o AQUAD (HUBER, 1991), entre otros. Pero quizá no sepamos que un procesador de textos, de las características de WORDPERFECT, 5.1 (WP51) puede ser tal útil e interesante como lo son otros creados específicamente para este tipo de trabajos. El procesador de textos WORDPERFECT

5.1. es un programa muy extendido y de gran utilidad entre los usuarios de la informática, al cual pueden tener acceso un mayor número de personas que a los programas anteriormente citados tan específicos y de difícil localización.

El trabajo que aquí se presenta pretende dar algunas pautas que puedan resultar de gran utilidad para aquellos interesados en el análisis de datos cualitativos, a la vez que usuarios del procesador de textos WP51.

EL ANÁLISIS DE DATOS

Antes de proceder a desarrollar algunas posibilidades de WP51 quizá convenga recordar, de manera esquemática, el proceso a seguir para llevar a cabo un análisis de datos una vez que los hemos recogido. Varios pueden ser los caminos para llegar a este proceso, pero acorde a criterios establecidos por autores de reconocida labor en este campo (BOGDAN & BIKLEN, 1982; MILES & HUBERMAN, 1984; PATTON, 1990), se sintetizan de la siguiente manera:

1. Preparar los datos

BOGDAN y BIKLEN (1982:145-155) insisten en la preparación de los datos recogidos como una función primordial a la hora de hacerlos frente para su análisis. Introducir en el ordenador la transcripción de las grabaciones, las notas de campo y observaciones recogidas en un formato de texto que mantenga unos criterios ordenados y estructurados para su sistematización posterior (preferiblemente en formato ASCII).

2. Codificar los datos

CRABTREE y MILLER (1992:93) organizan la información por medio de catálogos de códigos o «codebooks» con el fin de organizar el texto para su posterior interpretación. Establecer un sistema de códigos que permitan identificar el texto en unidades de significado. Cada código (de dos a ocho letras) representa una categoría diferente y excluyentes entre sí, de forma que nos permite reconocerla fácilmente entre las demás.

3. Organizar y reducir datos

Es el proceso de seleccionar, focalizar, simplificar y transformar los datos directos que han ido apareciendo en nuestro trabajo. La reducción de datos es una forma de analizar que afina, ordena, considera, descarta y organiza los datos de tal manera que puede obtenerse y verificarse las conclusiones «finales» (MILES & HUBERMAN, 1984:21).

UTILIDADES DE WP51 EN ESTE TIPO DE ANÁLISIS

Las utilidades que WP51 nos puede ofrecer a la hora de trabajar con los datos, entre otras, pueden ser las siguientes:

- * Numeración de líneas o párrafos de manera automática para su identificación en un texto escrito. (Ctrl+F8)
- * Búsqueda y sustitución de partes de texto o palabras. (Alt+F2)
- * Funciones «macro». (Alt+F10)
- * Documentos maestro y subdocumentos. (Alt+F5)

1. Preparar los datos

Inicialmente habremos de tener todos los documentos que queremos analizar en el ordenador, en ficheros de texto, perfectamente identificados y cada uno de ellos por separado. No es conveniente trabajar con ficheros excesivamente grandes por su tamaño en bytes. La configuración de la página de texto convendrá establecerla con un margen superior en el lado izquierdo. De esta manera podemos tener la numeración de líneas y la opción de ir colocando los códigos al lado de cada fragmento de texto que hayamos considerado como unidad de análisis (p.ej. párrafos de texto, en WP51 separados por dos [RtM]).

La opción de numeración de líneas en el texto escrito (ya que en pantalla no podemos visualizarla) se hace de la siguiente manera: Mayúsc+F8, (1 línea), (5 numeración de líneas).

Una vez que hemos configurado la página y sus márgenes (Mayúsc+F8) podemos imprimir el texto (Mayúsc+F7) con el fin de tener un original de los datos impresos que nos van a permitir establecer el siguiente paso de codificación manual de los datos.

2. Codificar los datos

Inicialmente el trabajo de codificar el texto, asignar códigos a las unidades de análisis, lo hemos realizado manualmente en el margen de la hoja que hemos preparado. A continuación, lo que vamos a desarrollar es el proceso de trasladar los códigos al texto que tenemos en el ordenador. Para ello vamos a proceder de la siguiente manera:

a) Recuperar el documento que queremos codificar (F5 lista de archivos), (1 Recuperar).

b) Con la opción de búsqueda (F2) de WP51 podemos localizar fácilmente aquella palabra o cadena de caracteres que queremos encontrar en el texto. Una vez localizado el lugar adecuado, colocamos al principio y al final de la cita o unidad codificada un signo, palabra o abreviatura, correspondiente al código utilizado para nuestro análisis. Este signo ha de ser distinto del resto del texto de forma que no pueda confundirse con una cadena de caracteres dentro de nuestros datos.

c) Una vez que tenemos hecha esta tarea utilizamos la función de búsqueda y sustitución (Alt+F2) que nos permite cambiar el signo o abreviatura por el código que queramos. De esta forma se agiliza mucho el trabajo de escribir códigos o palabras que identifican nuestro análisis, evitando tener que escribirla muchas veces.

d) Aprovechamos este momento para identificar todas aquellas unidades de análisis con una referencia al final de las mismas, de manera que cuando hagamos una agrupación de datos por códigos o categorías, tengamos perfectamente localizada a quien y a donde pertenece ese dato. Esta opción es cómoda de realizar a través de la función de búsqueda y sustitución (Alt+F2).

3. Agrupar y reducir datos

Este quizá sea el paso más decisivo en nuestro análisis con WP51. La forma como se va a desarrollar es a través de fines «macro». COVIELLA y COVIELLA (1990:605) definen la «macro» como una «combinación de teclas que automatizan tareas repetitivas, permitiendo un gran ahorro de tiempo en determinadas tareas».

La utilidad de las funciones macro nos van a permitir agrupar los datos en función de las categorías en las que los hemos asignado, pudiendo crear ficheros de texto de forma automática, sin tener que repetir la mecanografía de los mismos, a partir de la función de anexas textos (Ctrl+F4), (1 Mover bloque), (4 Anexar).

Los pasos que vamos a seguir en esta función son los siguientes:

- a) recuperar un fichero en el que tenemos los diferentes códigos asignados al texto.
- b) Crear una macro para cada código o categoría.

Para crear una macro para anexar cada código de texto vamos a seguir los siguientes pasos que a continuación se detallan:

Descripción del paso a seguir	Función en WP51
1. Mover el cursor al principio del documento.	Ini, Ini, ↑
2. Iniciar la definición de una macro.	Ctrl+F10
3. Dar nombre a la macro (p.ej. dar el nombre del código asignado PEC)	PEC
4. Describir la macro	Proyecto Educativo de Centro
5. Buscar los caracteres utilizados como código	F2 (buscar: &PEC)
6. Activar la opción de bloque	Alt+F4
7. Buscar hasta el final de la unidad asignada como categoría codificada (la unidad de análisis es el párrafo, separado por dos [RtM])	F2 (buscar &PEC[RtM][RtM])
8. Estando el bloque activado, seleccionar la función de añadir el bloque de texto a un fichero como el lugar donde agrupamos los datos que tienen el mismo código.	Ctrl+F4, 1, 4
9. Damos el nombre al fichero donde queremos guardar estos datos	Anexar en (p.ej. PEC.COD)
10. Invocamos la macro que acabamos de crear, de forma que se siga realizando de forma automática hasta el final del texto.	Alt+F10 (ej. PEC)
11. Terminamos de definir la macro.	Ctrl+F10

Estos pasos habremos de realizarlos para cada uno de los códigos establecidos en nuestro trabajo.

Una vez que ya hemos hecho las macros individuales para cada código, ahora sólo nos queda trabajar con todos los documentos de forma automática, haciendo que WP51 haga esta tarea de agrupar el texto con los mismas categorías en ficheros independientes.

Anteriormente se dijo que no conviene trabajar con ficheros de gran tamaño, pero ahora vamos a crear un documento maestro que contenga todos los subdocumentos o ficheros.

Con la función de subdocumento (Alt+F5) podemos poner uno detrás de otro y tener en un documento maestro todos los ficheros de texto. Una vez que tenemos todos los documentos que contienen nuestros datos, deberemos expandir el documento maestro a través de la orden Alt+F5, 6 (Generar), 3 (Expandir el documento maestro).

Cuando tengamos expandido en pantalla todos los datos, nos colocamos al principio del documento (Ini, Ini, ↑) y activamos la función macro que hemos definido anteriormente (Alt+F10, nombre: PEC). El proceso que sigue a continuación es que WP51 va a agrupar en un fichero todas las unidades de texto que hemos codificado con el mismo criterio.

Una vez que ha terminado este proceso, volvemos al principio del documento (Ini, Ini, ↑) y activamos otra macro (Alt+F10), correspondiente a otro código, y así sucesivamente. El resultado final será que podemos obtener documentos de texto en los que están agrupadas las categorías asignadas a diferentes documentos. Esto nos ayudará a considerar nuestro trabajo como combinación de funciones que WP51 nos brinda como procesador de texto de gran potencia y versatilidad.

REFERENCIAS BIBLIOGRÁFICAS

- BOGDAN, R. C. & BIKLEN, S. K. (1982): *Qualitative Research for Education*. London: Allyn & Bacon.
- BRACKMAN, B. J. (1987): *Qualpro Text Database and Productivity, (User's Manual)*. Tallahase, FL.: Impulse Development Company.
- CRABTREE, B. F. & MILLER, W. L. (1992): A Template Approach to Text Analysis: developing and Using Codebooks. In B. F. CRABTREE & W. L. MILLER (Eds.): *Doing Qualitative Research*. Newbury Park, CA.: Sage, pp. 93-109.
- COVIELLA, J. M. & COVIELLA, E. (1990): *Trabajando con WordPerfect 5.1*. Madrid: Ra-Ma.
- DRASS, K. A. (1980): The Analysis of Qualitative Data: A Computer Program. *Urban Life*, 9 (3): 332-353.
- HUBER, G. (1991): *Análisis de datos cualitativos con ordenadores, Principios y manual del paquete de programas AQUAD 3.0*. Sevilla: Grupo de Investigación Didáctica.
- MANSFIELD SOFTWARE GROUP (1987): *Kedit User's Manual, Version 3.52*. Storrs, CT.: Author.
- MILES, M. B. & HUBERMAN, A. M. (1984): *Qualitative Data Analysis*. Beverly Hills, CA.: Sage.
- PATTON, M. Q. (1990): *Qualitative Evaluation and Research Methods*. 2nd. Edition. Newbury Park, CA: Sage.
- PFAFFENBERGER, B. (1988): *Microcomputer Applications in Qualitative Research*. Newbury Park, CA.: Sage.
- REID, A. O. (1992): Computer Management Strategies for Text Data. In B. F. CRABTREE & W. L. MILLER (Eds.): *Doing Qualitative Research*. Newbury Park, CA.: Sage, pp. 125-145.
- SEIDEL, I. V., KJOLSETH, R. & SEYMOUR, R. (1988): *The Ethnograph. A User's Guide (Version 3.0)*. Littleton, CO.: Qualis Research Associates.

ANÁLISIS ESTADÍSTICO DE DATOS CUALITATIVOS TEXTUALES: EL ENFOQUE LEXICOMÉTRICO

Javier Gil, Eduardo García, Gregorio Rodríguez y Ana Corrales
Universidad de Sevilla

El propósito de este trabajo es revisar el modo en que es analizada la información textual desde la perspectiva lexicométrica desarrollada por la escuela francesa de análisis de datos, con el fin de hacer una valoración acerca de los procedimientos incluidos bajo este enfoque.

1. ANÁLISIS CUANTITATIVO DE DATOS CUALITATIVOS

Con frecuencia, el investigador educativo se encuentra ante datos de naturaleza textual, es decir, datos expresados en forma de cadenas verbales de cierta extensión y que generalmente acaban siendo registrados como textos escritos. El análisis cualitativo de este tipo de datos se ha realizado a menudo a través de procesos de naturaleza intuitiva y creativa, que han dado pie a las críticas de quienes caracterizan el análisis de los datos textuales puesto en práctica por los investigadores cualitativos como oscuro, difuso, falto de rigor, poco definido.

A pesar de que son cada vez más habituales los esfuerzos por afrontar la indefinición de los procedimientos y pautas por los que se lleva a cabo el análisis, tratando de clarificar, guiar, dar rigor y exhaustividad a las técnicas analíticas cualitativas (Taylor y Bogdan, 1986; Lofland y Lofland, 1984; Miles y Huberman, 1984), son muchos los investigadores que asumen una perspectiva cuantitativa al analizar los datos cualitativos textuales (Berelson, 1952), recurriendo en algún momento de su proceso analítico a las técnicas estadísticas, con la intención de complementar o contrastar las conclusiones obtenidas por otras vías. Incluso este rasgo ha sido visto como característico de los enfoques de investigación interpretativos o cualitativos, los cuales recurren, sobre todo en el análisis de datos, a procedimientos cuantitativos característicos del enfoque positivista (De Miguel, 1988). Las ventajas del análisis de datos numéricos ha contribuido a un aumento en el uso del análisis estadístico en la investigación cualitativa, que ha quedado reflejado en la literatura a través de un incremento sostenido, durante las últimas décadas, de los trabajos que incluyen análisis estadísticos (Pelto y Pelto, 1991).

Parece evidente que, aunque las palabras aportan mayor significado, los números resultan menos ambiguos y son procesados con mayor economía. Por esta razón, buena parte de los investigadores prefiere trabajar con números y trasladan los textos a números lo más rápidamente posible. Generalmente, los valores numéricos surgen a partir del cómputo de los elementos diferenciados —frecuencias de códigos empleados en la categorización— al analizar el corpus de datos cualitativos textuales. Para

los investigadores que toman este camino, el terreno está bien marcado por las teorías de la medición, reglas de decisión, niveles de confianza, algoritmos para el cálculo, etc. (Miles y Huberman, 1984).

La cuantificación y el análisis estadístico son, por tanto, herramientas analíticas con las que cuenta el investigador en su trabajo con datos cualitativos, y pueden ser utilizadas conjuntamente con otras herramientas no cuantitativas. La formación del investigador, sus objetivos y sus concepciones epistemológicas le llevan a configurar un proceso de análisis orientado en mayor o menor medida hacia uno de estos tipos de técnicas. Dependiendo del estudio, la extensión de la cuantificación de datos varía ampliamente, desde ninguna hasta el uso de técnicas estadísticas de variado grado de complejidad (Wilcox, 1982). La cuantificación puede llegar a ser el aspecto central del análisis, e incluso a veces los datos cualitativos son, desde el primer momento, trasladados a índices numéricos y analizados cuantitativamente. En esta línea se inscriben los enfoques cuantitativos lexicométricos, apoyados en las técnicas estadísticas desarrolladas por la escuela francesa de análisis de datos.

2. EL ENFOQUE LEXICOMÉTRICO EN EL ANÁLISIS DE TEXTOS

Lexicometría es el término bajo el que se agrupan una serie de métodos que permiten reorganizar las unidades presentes en una secuencia textual y operar ciertos análisis estadísticos a partir de la cuantificación del vocabulario resultante de una segmentación operada sobre el texto (Lebart y Salem, 1988). Las posibilidades actuales en el campo de la lexicometría se han visto ampliamente incrementadas con el uso del ordenador; los primitivos análisis basados en la construcción de glosarios de términos y en el estudio de las concordancias (diferentes contextos verbales en los que ocurre una misma palabra) han pasado a constituir el material de partida para la aplicación de métodos estadísticos multidimensionales, basados en las técnicas de análisis de datos desarrolladas por Bénzecri (1973) —métodos factoriales y de clasificación, fundamentalmente—, que permiten extraer conclusiones acerca del sentido de un texto.

La aplicación de las técnicas del análisis de datos de la escuela francesa a datos de naturaleza textual (respuestas a cuestionarios de preguntas abiertas, entrevistas, textos literarios...etc.), apoyándose en las posibilidades ofrecidas por el ordenador, representa un intento de eludir la precodificación del texto y trabajar directamente con el corpus textual. De esta forma, resulta posible estudiar el sentido latente del texto a partir de los datos brutos, evitando el riesgo de la inconsistencia en la codificación o la pérdida de información que comporta la mediación de una categorización inicial.

Los estudios lexicométricos surgen, por tanto, de la necesidad sentida en el estudio de los textos de sobrepasar los enfoques tradicionales, marcados por la subjetividad, y abordar su análisis a partir del recuento y la localización de las unidades que contienen (Salem, 1982), considerándose como tales la palabra o una secuencia de palabras. La unidad considerada como base de los recuentos es generalmente la **palabra**, como ha ocurrido en otros enfoques del análisis de contenido de carácter cuantitativo realizado con ordenador (Mochmann, 1985; Mohler, 1985). La palabra cuenta con la ventaja de ser una unidad formal fácilmente identificable por el ordenador, al estar separada de otras unidades mediante un espacio o un signo de puntuación (Sánchez Carrión, 1985), y ello la hace especialmente adecuada para tratamientos automáticos. Lebart y Salem (1988), denominan *forma gráfica* a la unidad básica en que dividen a la cadena textual, y la definen del siguiente modo:

«La unidad básica de los recuentos será la forma gráfica, definida como una sucesión de caracteres no delimitadores (en general, letras) rodeada por dos caracteres delimitadores (espacios, puntos, comas,...)» (p. 6).

Un riesgo implícito en el uso de formas gráficas como unidades radica en que diferentes palabras,

a veces con la misma raíz léxica, pueden expresar significados equivalentes, o podemos encontrarlos también ante palabras polisémicas, con más de un significado. Ambos aspectos deben ser tenidos en cuenta a efectos de enumeraciones posteriores. Por esta razón, Bolasco (1992:70) propone un tipo alternativo de unidad: la *forma textual*, identificada con una raíz léxica que corresponde a una sola forma gráfica o a sus diferentes inflexiones, siempre que sean portadoras de un significado indudablemente equivalente para las finalidades del estudio, o identificada con diferentes raíces léxicas que corresponden a diferentes formas gráficas con igual significado. Sin embargo, trabajar con formas textuales implica una intervención previa del analista, que debe definir las equivalencias entre formas gráficas o reducirlas a raíces léxicas comunes (lematización).

La unidad para el análisis puede ser un **grupo de palabras**. Lebart y Salem (1988), considerando que descender hasta el nivel de palabras puede conducir a un aislamiento que haga perder el sentido de lo expresado en el texto, utilizan también como unidad los *segmentos repetidos*, constituidos por dos o más palabras delimitadas por separadores tales como los signos de puntuación.

A partir de la segmentación, el corpus textual queda transformado en *tablas lexicales*, en las que se distribuyen las unidades para individuos o conjuntos de individuos con determinadas características. Estas tablas recogen ordenadamente la información sobre ocurrencia de las formas gráficas, situando en las columnas las partes consideradas en el texto (sobre la base de momentos diferentes, emisores distintos, etc.) y en las filas las formas del corpus textual o sólo aquéllas que rebasan una determinada frecuencia. Los términos son colocados en orden lexicométrico: frecuencias o longitudes decrecientes y/u orden alfabético. La aplicación de procedimientos estadísticos multivariantes de carácter descriptivo, exploratorio, a este tipo de tablas arroja como resultado índices, agrupamientos, representaciones gráficas que permiten profundizar en la comprensión de los datos y de la forma en que se estructuran. Del mismo modo, es posible realizar distintos análisis a partir de *tablas de segmentos repetidos*.

Las técnicas del análisis de datos de la escuela francesa son especialmente adecuadas para el tratamiento de valores numéricos procedentes de datos de naturaleza cualitativa. Precisamente, el tratamiento cuantitativo de lo cualitativo y el análisis simultáneo de un conjunto de variables definen las aportaciones fundamentales de tales métodos (Cornejo, 1988). Se trata de métodos, a la vez exploratorios y multidimensionales, que permiten descubrir la estructura de una gran tabla de números de varias dimensiones, traduciéndola a otra estructura más simple y que la resume mejor (Clapier, 1983).

3. EL PAQUETE ESTADÍSTICO SPAD.T

Dentro del enfoque lexicométrico desarrollado por la escuela francesa de análisis de datos (Lebart Salem, 1988), se ha diseñado un software específico capaz de trabajar con datos textuales procedentes de respuestas abiertas a cuestionarios, entrevistas, textos literarios, etc. El programa SPAD.T —Système Portable pour l'Analyse des Données Textuelles— (Lebart, Morineau y Bécue, 1989) permite extraer, de este tipo de datos, tablas de frecuencias construidas sobre las palabras o grupos de palabras presentes en una serie de textos para llevar a cabo diversos análisis lexicométricos, a los que añade métodos estadísticos como el análisis de correspondencias, la clasificación automática y el cálculo de las especificidades.

El **análisis de correspondencias**, en este caso, básicamente consiste en transformar la información de una tabla lexical en un conjunto de puntos, correspondientes a individuos o variables, posicionados en un espacio multidimensional, que pueden ser proyectados en los planos que mejor permiten visualizar la estructuración del conjunto. Las semejanzas y diferencias entre formas gráficas y/o puntos-columna quedan unívocamente reflejadas por las distancias entre los puntos que los representan en el espacio geométrico considerado. La interpretación de los gráficos factoriales resultantes consistirá en extraer conclusiones a partir del juego de proximidades y oposiciones puestas de relieve (Cibois, 1991).

Los **métodos de clasificación** permiten agrupar en clases, o jerarquías de clases, objetos o individuos sobre los cuales poseemos cierta información. Utilizan procedimientos basados en cálculos algorítmicos y producen representaciones gráficas generalmente en forma de árbol invertido.

El **cálculo de las especificidades** o formas características de los textos (Lafon, 1980) puede resultar de gran ayuda en la interpretación de las correspondencias reflejadas en los planos factoriales. Permite establecer qué formas se distribuyen uniformemente en el corpus de datos y cuáles no lo están y son características de una de las partes consideradas en éste. Se apoya en la obtención de un valor estadístico basado en la comparación de los porcentajes de frecuencia global e interno para cada forma gráfica. A cada forma se asocia un valor de contraste cuya magnitud aumenta a medida que disminuye la probabilidad de encontrar una determinada subfrecuencia interna. Este estadístico presenta una distribución normal. Por tanto, sus valores pueden ser trasladados a una curva normal para conocer la probabilidad asociada a la hipótesis nula que postula la repartición aleatoria de la forma entre los textos.

4. VALORACIÓN DEL ENFOQUE LEXICOMÉTRICO

El análisis lexicométrico permite enfrentarse a los datos textuales eludiendo la codificación de los mismos, y por tanto, el componente subjetivo, el tiempo y el esfuerzo asociados a esta operación. Se consigue así dar respuesta satisfactoria a la aspiración compartida por una buena parte de los investigadores de desarrollar el análisis de los datos de modo que cualquier analista siguiendo el mismo proceso llegue a los mismos resultados, objetivo que no resulta posible cuando se introducen en el proceso fases de codificación de los datos, incluso aunque se hagan esfuerzos por seguir un mismo patrón o esquema.

Considerado como lectura automática de textos, el análisis lexicométrico permite aproximarnos a la estructura de la información contenida en los textos, planteando interrogantes, abriendo vías de exploración y, en definitiva, contribuyendo eficazmente al planteamiento de hipótesis en las fases iniciales del análisis. Al apoyarse en técnicas descriptivas, resulta idóneo para ser aplicado cuando no existen hipótesis previas de trabajo y se requiere una exploración de la realidad estudiada con el fin de establecer los puntos de partida para posteriores análisis (García Santesmases, 1984).

Sin embargo, este tipo de procedimientos que recurren a la cuantificación es frecuentemente objeto de controversia, dado que no todos los investigadores están de acuerdo en considerar que los números puedan dar cuenta de las estructuras de significado contenidas en los datos textuales, sin que ello implique una pérdida excesiva de la riqueza informativa que les caracteriza. La segmentación del texto implica la pérdida de los significados contextuales de las palabras, limitando el estudio a la «superficie discursiva» del texto e ignorando el fondo significativo que subyace a ella. Implica un tratamiento indiferenciado de las palabras con una misma forma, que presupone una relación biunívoca entre el sentido con que son empleados los significantes objeto de recuento y sus significados, ignorando fenómenos como la polisemia y homonimia. Un problema inverso surge con la dispersión de formas que aluden a un mismo significado. Es el caso de los sinónimos y, con mayor amplitud, de las múltiples inflexiones de un verbo.

Se trata de un tipo de análisis subsidiario de las aplicaciones informáticas. Sin ellas, teniendo en cuenta que la unidad considerada en el análisis es la palabra, la tarea de extraer, ordenar y proceder al recuento de las apariciones de cada una de los vocablos presentes en un texto sería una tarea prácticamente inabordable cuando trabajamos con textos de cierta extensión. Del mismo modo, las técnicas estadísticas a que son sometidas las frecuencias precisan del soporte informático si queremos operar sobre grandes masas de datos. El programa SPAD.T reúne las características apropiadas para abordar este tipo de análisis.

A pesar del recelo que en algunos investigadores pudieran despertar el uso de la cuantificación o la rigidez del análisis automatizado, consideramos que la información textual puede ser analizada por

técnicas estadísticas como las representadas por el enfoque lexicométrico. La primera aproximación al significado de los datos que tales técnicas nos proporcionan, además de descubrirnos diferencias entre distintas partes del corpus textual, significados latentes, estructuras en los datos,..., resulta especialmente útil en la formulación de hipótesis o en el establecimiento de categorías para posteriores análisis cualitativos.

REFERENCIAS BIBLIOGRÁFICAS

- BENZECRI, J. P. (1973): *L'Analyse des Données*. París: Dunod.
- BERELSON, B. (1952): *Content Analysis in Communications Research*. Nueva York: Free Press.
- BOLASCO, S. (1992): Sur différentes stratégies dans une analyse des formes textuelles: une expérimentation à partir de données d'enquête. En M. Becue, L. Lebart y N. Rajadell: *Jornades Internacionals d'Anàlisi de Dades Textuals*. Barcelona: Servicio de Publicaciones de la UPC. (pp. 69-88).
- CIBOIS, PH. (1991): *L'analyse factorielle*. París: Presses Universitaires de France.
- CORNEJO, J. M. (1988): *Técnicas de investigación social: El análisis de correspondencias*. Barcelona: PPU.
- CLAPIER, P. (1983): *Análisis de Datos*. Vitoria: Publicaciones del Gobierno Vasco.
- DE MIGUEL, M. (1988): Paradigmas de la Investigación Educativa Española. En I. Dendaluce (Coord.). *Aspectos Metodológicos de la Investigación Educativa. II Congreso Mundial Vasco*. Madrid: Narcea, (pp. 60-77).
- GARCÍA SANTESMASES, J. M. (1984): Análisis factorial de correspondencias. En J. J. Sánchez Carrión, *Introducción a las técnicas de análisis multivariable aplicadas a las ciencias sociales*. Madrid: Centro de Investigaciones Sociológicas, (pp. 75-105).
- LAFON, P. (1980): Sur la variabilité de la fréquence des formes dans un corpus. *Mots*, (1), 127-165.
- LOFLAND, J. y LOFLAND, L. H. (1984): *Analyzing social settings. A guide to qualitative observation and analysis*. Belmont, CA: Wadsworth.
- LEBART, L. y SALEM, A. (1988): *Analyse Statistique des Données Textuelles. Questions ouvertes et Lexicométrie*. París: Bordas.
- LEBART, L., MORINEAU, A. y BECUE, M. (1989): *SPAD-T. Système portable pour l'analyse des données textuelles. Manuel de l'utilisateur*. París: CISIA.
- MILES, M. B. y HUBERMAN, A. M. (1984): *Qualitative Data Analysis. A Sourcebook of New Methods*. Beverly Hills: Sage Publications.
- MOCHMANN, E. (1985): Análisis de Contenido mediante Ordenador Aplicado a las Ciencias Sociales. *Revista Internacional de Sociología*, 43 (1), 11-44.
- MOHLER, P. Ph. (1985): Algunas Observaciones Prácticas sobre la Utilización del Análisis de Contenido en Ordenador. *Revista Internacional de Sociología*, 43 (1), 45-57.
- PELTO, P. J. y PELTO, G. H. (1991): *Anthropological research. The structure of inquiry*. Nueva York: Cambridge University Press.
- SALEM, A. (1982): Analyse factorielle et lexicométrie. Synthèse de quelques expériences. *Mots*, (4), 147-168.
- SÁNCHEZ CARRIÓN, J. J. (1985): Técnicas de Análisis de los Textos mediante Codificación Manual. *Revista Internacional de Sociología*, 43 (1), 89-118.
- TAYLOR, S. J. y BOGDAN, R. (1988): *Introducción a los métodos cualitativos de investigación*. Buenos Aires: Paidós.
- WILCOX, K. (1982): Ethnography as a methodology and its application to the study of schooling: a review. En G. Spindler (Ed.): *Doing the Ethnography of Schooling: Educational anthropology in action*. Nueva York: Holt, Rinehart and Winston, (pp. 456-488).

FORMULACIÓN Y COMPROBACIÓN DE HIPÓTESIS CUALITATIVAS CON AYUDA DEL PROGRAMA «AQUAD»

Javier Gil, Eduardo García, Gregorio Rodríguez y Ana Corrales
Universidad de Sevilla

Las operaciones que normalmente se asocian a los procesos de análisis de datos cualitativos (García et al., 1993) se han visto enormemente facilitadas por la incorporación del ordenador, cuyas posibilidades de manejo de gran cantidad de datos numéricos, verbales o visuales, y su rapidez en la realización de cálculos o tareas de búsqueda y organización le han situado en un lugar de privilegio también en el ámbito de la investigación educativa y, en particular, en el análisis de datos cualitativos textuales. Las aplicaciones informáticas han venido a reducir la enorme cantidad de esfuerzo y tiempo que consumen las tareas de análisis de este tipo de datos. Las ventajas que la utilización de procesadores de texto o bases de datos ofrecen —permitiendo la entrada y el almacenamiento de los datos textuales, la ordenación de una serie de ítems o la búsqueda de informaciones específicas— han sido superadas por programas específicamente diseñados para el análisis del material cualitativo. Este tipo de programas se inscriben en las líneas cualitativas de investigación, y su contribución es importante de cara a realizar el proceso de análisis de datos cualitativos de un modo que pueda ser controlado, reconstruido y comunicado.

En este trabajo, hemos seleccionado el programa AQUAD —Analysis of Qualitative Data— (Huber, 1988) que, además de facilitar las funciones habitualmente recogidas en el software diseñado para esta tarea, tales como la codificación, recuperación de unidades o el recuento de ocurrencias y coocurrencias, permite formular y comprobar condiciones lógicas subyacentes a potenciales relaciones entre los códigos. Nuestra atención se dirigirá especialmente al estudio de sus posibilidades para la formulación y comprobación de hipótesis en el marco de la investigación cualitativa, entendiendo por tales hipótesis *relaciones entre tópicos identificados en un texto a partir de las cuales es posible reconstruir un todo estructurado que nos aproxima a una comprensión global de los datos.*

EL ANÁLISIS MEDIANTE AQUAD

El programa AQUAD ofrece un amplio abanico de posibilidades a los investigadores que se enfrentan a la tarea de analizar datos cualitativos, contribuyendo en las fases de reducción de datos, presentación de los mismos y extracción de conclusiones (Huber y Marcelo, 1990). Aunque aquí nos interesa especialmente la fase de comprobación de hipótesis (fase con la que AQUAD se distancia de otros programas para el análisis de datos cualitativos), no podemos perder de vista el proceso completo de análisis que permite desarrollar este programa. El procedimiento de análisis que aquí presentamos queda esquematizado en la Figura 1.

1. Codificación de los textos

Una operación habitual en el análisis de datos cualitativos es la reducción de datos, que suele concretarse en la codificación de los datos textuales. AQUAD permite al analista señalar, sobre el texto visualizado en pantalla, el inicio y el final de cada fragmento considerado como una unidad de análisis y etiquetarlo con un código. La codificación puede ser múltiple: un mismo fragmento puede ser codificado por varios códigos, los textos etiquetados con códigos pueden superponerse parcialmente, o quedar incluidos unos dentro de otros.

Durante la codificación puede resultar útil revisar el contenido de las categorías, examinando el contenido de los fragmentos incluidos en cada una de ellas con objeto de asegurarnos de la consistencia en la codificación. La recuperación de fragmentos etiquetados con un determinado código podría realizarse de forma automática mediante AQUAD. En cualquier caso, el ordenador cumple tareas mecánicas, como la búsqueda de fragmentos codificados, pero no conceptuales; no puede sustituir al investigador en la toma de decisiones sobre la asignación de tales fragmentos a determinadas categorías.

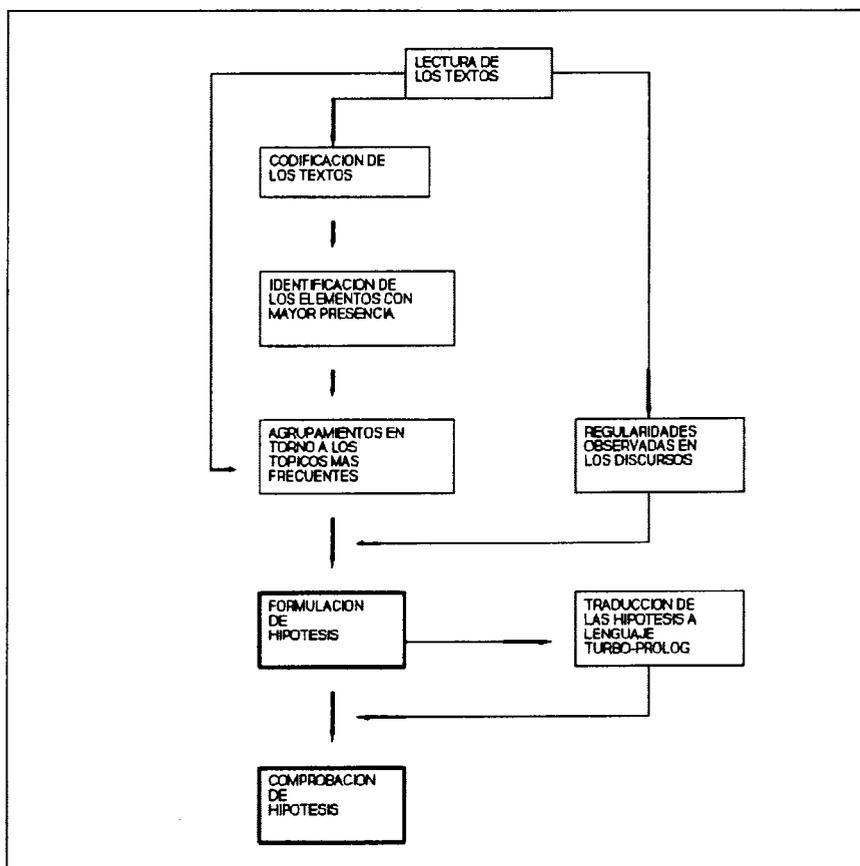


Figura 1. Proceso seguido en el análisis de los datos usando el programa AQUAD.

2. Identificación de los elementos con mayor presencia

A partir de la codificación, podemos identificar los temas que aparecen con mayor frecuencia en el conjunto de textos codificados. Tales tópicos serán los más relevantes de cara al tema investigado o, cuando menos, los aspectos sobre los cuales se posee mayor información. Puesto que nuestras conclusiones se basarán en la presentación de evidencia que apoye ciertas afirmaciones, es necesario centrar la atención sobre los tópicos de mayor presencia en los textos. Identificar los códigos más frecuentes es una operación bastante fácil de realizar mediante AQUAD, que puede facilitarnos un recuento de las apariciones de cada código en cada uno de los textos sometidos a análisis.

3. Agrupamientos en torno a los códigos más frecuentes

Identificados los tópicos más frecuentes, es necesario buscar qué otros tópicos se encuentran relacionados con éstos. Un primer paso para alcanzar tal objetivo consiste en analizar el agrupamiento físico de códigos que se da en torno a ellos. Estudiar los agrupamientos consistiría en localizar códigos cuyos fragmentos asociados coaparecen en el texto junto a los de los códigos más frecuentes. Como criterio de agrupamiento podríamos adoptar el no distanciarse en más de un determinado número de líneas del código considerado. Es decir, entenderíamos que se agrupan en torno a un código aquellos códigos cuyas áreas de texto correspondientes anteceden, se superponen o suceden al área de texto de un código determinado, sin que entre ambas áreas exista una distancia, en número de líneas, superior a un valor fijado.

En lugar de revisar una gran cantidad de páginas buscando tales agrupamientos, AQUAD podría facilitarnos este tipo de coapariciones, a partir de las cuales, como se verá en el próximo apartado, se intuyen regularidades y estructuras de significado conducentes a la formulación de hipótesis.

4. Formulación de hipótesis

Para entender la naturaleza de la posible relación entre los códigos que coaparecen en mayor número de ocasiones, es preciso volver a los textos codificados a fin de revisar algunos de los fragmentos correspondientes. Esta fase del proceso supone, por tanto, una vuelta a los datos para establecer si la coaparición de los códigos obedece a una simple yuxtaposición temática, a una secuenciación de tópicos, a una relación causa-efecto, a una inclusión de determinados temas en otros, a una contraposición de elementos, etc. A la formulación de hipótesis también contribuyen las ideas que pudieran haber surgido en el analista a partir de la lectura inicial de los textos.

Las hipótesis formuladas se refieren a las posibles regularidades de patrones de significado —constelaciones de códigos— presentes en los datos; plantean la existencia de una relación entre dos o más códigos, entendiendo que esta relación se traduce en proximidad espacial dentro del texto, inclusión entre unidades o secuencia de aparición. De este modo, el programa permite extraer conclusiones respecto a configuraciones de unidades de significado que se repiten a lo largo de un corpus de datos, cuya comprobación manual resultaría una tarea de complejidad considerable.

Para ilustrar el modo en que AQUAD contribuye a la formulación de hipótesis cualitativas, traemos aquí un ejemplo de hipótesis en la que se establece una relación entre dos códigos. Los datos utilizados corresponden a una investigación reciente en la que se pedía a grupos de profesores que opinaran sobre la reforma educativa (Gil Flores, 1992). Los datos textuales producidos en dos de estos grupos han sido utilizados para ejemplificar el proceso descrito en el presente trabajo.

Codificado el texto y realizado el recuento de códigos, hemos destacado el código «COP» (consecuencias de la aplicación de la reforma para el profesorado) como uno de los más frecuentes. El

posterior examen de agrupamientos en torno a este código, revela que el código «STR» (aspectos estructurales de la reforma) es el que en mayor medida coaparece junto a éste. Para determinar el tipo de relación que pudiera existir entre ambos, revisamos algunos de los fragmentos de texto implicados. Así, por ejemplo, citamos aquí una de las coocurrencias de los códigos COP y STR en nuestro corpus de datos, localizada con la ayuda del programa AQUAD:

1254 Yo, por lo que conozco de la LOGSE, el ^STR-1263
1255 diseño de la Primaria yo lo encuentro
1256 bien. Ahora, para mí, el meollo de la
1257 reforma es precisamente la Secundaria. A
1258 mí me parece... yo veo unas lagunas tan
1259 gordas, que yo es que no entiendo cómo se
1260 va a implantar la Secundaria. Es que no lo
1261 sé. Quién va a dar el primer ciclo, quién ^COP-1263
1262 el segundo, los licenciados, los
1263 profesores de la Segunda Etapa.

La lectura de las áreas de texto referidas en nuestro ejemplo permite establecer que las consecuencias para el profesorado se derivan de la nueva estructuración del sistema educativo. Esto concuerda con la impresión obtenida durante la lectura global de los textos, en la cual advertíamos una cierta preocupación por el modo en que se verán afectados los profesores por los aspectos estructurales, pero no por los curriculares (los fragmentos en los que se aludía a aspectos curriculares fueron codificados con CUR).

Consecuentemente, la hipótesis planteada afirmaría que *los profesores contemplan las consecuencias que tendrá para ellos la aplicación de la reforma como una derivación de los aspectos estructurales y no de los curriculares.*

5. Comprobación de hipótesis

Las hipótesis serán finalmente contrastadas en el conjunto de los datos. Debe tenerse en cuenta que esta comprobación no es un contraste de hipótesis en sentido teórico, puesto que no existe una distribución teórica de referencia a la que comparar los resultados obtenidos para un determinado estadístico ni existen criterios fijados para la aceptación o el rechazo de las mismas. Se trata de un contraste en sentido empírico, consistente en explorar la frecuencia con que determinadas «constelaciones» de códigos se dan en el corpus de datos, indicando tendencias que se ven, de este modo, apoyadas por un mayor o menor número de ocurrencias dentro del mismo. En este sentido, en lugar de formulación y comprobación de hipótesis, quizás sería más adecuado hablar de relaciones entre códigos y búsqueda de evidencia para apoyar la existencia de tales relaciones.

Para que las hipótesis puedan ser comprobadas con la ayuda de AQUAD, las relaciones que establecen deben ser trasladadas al lenguaje de programación turbo-prolog, de modo que el ordenador proceda sistemáticamente a comparar todos los elementos de significado presentes en el conjunto de textos, de acuerdo con condiciones lógicas especificadas para ellos. Retomando el ejemplo que presentábamos en el apartado anterior, para comprobar la hipótesis formulada en nuestro corpus de datos, habría que examinar si los códigos COP van precedidos por códigos STR, pero no por códigos CUR. Entenderemos que COP está precedido por STR o CUR cuando el fragmento correspondiente a COP no empieza antes que el de STR o CUR ni a una distancia mayor de 5 líneas después de su final. La comprobación de estas condiciones en cada pareja formada por un código COP y un código CUR

o STR constituye el eje del programa en turbo-prolog que permite comprobar la hipótesis de nuestro ejemplo. El programa completo quedaría del siguiente modo:

```

hy(1,1):-write("APARICIONES DEL CODIGO COP","\n").
hy(1,1):-C="COP", k(_,A1,E1,C), write(A1,E1,C,"\n"),fail.
hy(1,1):-write("ASPECTOS ESTRUCTURALES PRECEDIENDO A COP:", "\n").
hy(1,1):-C1="COP",C2="STR",D=5,
    k(N,A1,E1,C1), str_int(A1,IA1),
    k(N,A2,E2,C2), str_int(A2,IA2),str_int(E2,IE2),
    IA1-IA2>=0,IA1-IE2<=D,
    write(A2,E2,C2," - ",A1,E1,C1,"\n"),fail.
hy(1,1):-write("ASPECTOS CURRICULARES PRECEDIENDO A COP:", "\n").
hy(1,1):-C1="COP",C2="CUR",D=5,
    k(N,A1,E1,C1), str_int(A1,IA1),
    k(N,A2,E2,C2), str_int(A2,IA2),str_int(E2,IE2),
    IA1-IA2>=0,IA1-IE2<=D,
    write(A2,E2,C2," - ",A1,E1,C1,"\n"),fail.

```

Las instrucciones presentadas aquí y escritas al final del fichero AQD-H.PRO, se ejecutarían sobre los ficheros de códigos creados por AQUAD para cada texto codificado. En tales ficheros, cada código presenta la estructura $k(N,A,E,C)$, donde N es la extensión con que se nombra el fichero de texto a analizar, A y E son las líneas en que comienza y termina el área marcada por un código, y C la etiqueta de tres caracteres que identifica al código.

De acuerdo con lo formulado, en las dos primeras líneas se indica al ordenador que busque los códigos COP y escriba las líneas de comienzo y final de los párrafos correspondientes. El siguiente bloque de instrucciones comienza asignando valores a las variables código (C1 y C2) y distancia (D). Las líneas quinta y sexta implican la búsqueda de una pareja de códigos C1 y C2 y la transformación en números enteros (IA1, IA2, IE2), con las funciones «str_int», de los parámetros que indican principio y final de un área de código (A1, A2, E2), expresados inicialmente como cadenas de caracteres. Los números obtenidos deberán cumplir las condiciones establecidas en la línea siguiente, es decir, de acuerdo con la condición $IA1-IA2_0$, al restar la línea en que comienza el fragmento de STR a la línea en que comienza COP deberíamos obtener 0 ó un número positivo. Ello significaría que el número IA1 es mayor o igual que IA2, y por tanto, el fragmento de COP comienza después o al mismo tiempo que el de STR. Por otro lado, $IA1-IE2_5$ impone la condición de que la diferencia entre el inicio de COP y el final de STR no supere las 5 líneas, o lo que es igual, que COP no comience a una distancia superior a las 5 líneas después del final de STR. El programa escribirá los parámetros que definen a los fragmentos que cumplan tales condiciones, de acuerdo con lo especificado en la séptima línea. La instrucción «fail» indica que se vuelva a desarrollar el proceso hasta que hayan sido considerados todos los casos y soluciones posibles. El proceso seguido con la pareja de códigos COP y STR sería repetido con COP y CUR, como indican las restantes líneas del programa.

El resultado de ejecutar este programa es la localización en el texto de las parejas de códigos que presentan la estructura exigida por la hipótesis, para que en función de la evidencia empírica podamos valorar la plausibilidad de ésta. En el ejemplo que desarrollamos, los resultados fueron:

***** textouno.cod

APARICIONES DEL CODIGO COP

110	126	COP
142	155	COP
252	269	COP
289	307	COP
423	424	COP
694	699	COP
1586	1592	COP
1837	1845	COP
1860	1861	COP

ASPECTOS ESTRUCTURALES PRECEDIENDO A COP:

110	118	STR	-	110	126	COP
252	269	STR	-	252	269	COP
289	293	STR	-	289	307	COP
421	422	STR	-	423	424	COP
675	699	STR	-	694	699	COP
1832	1838	STR	-	1837	1845	COP
1857	1861	STR	-	1860	1861	COP

ASPECTOS CURRICULARES PRECEDIENDO A COP:

1586	1592	CUR	-	1586	1592	COP
------	------	-----	---	------	------	-----

***** textodos.cod

APARICIONES DEL CODIGO COP

36	38	COP
1156	1163	COP
1184	1189	COP
1230	1234	COP
1240	1244	COP
1261	1263	COP
1505	1507	COP
1732	1746	COP

ASPECTOS ESTRUCTURALES PRECEDIENDO A COP:

1139	1156	STR	-	1156	1163	COP
1168	1184	STR	-	1184	1189	COP
1218	1234	STR	-	1230	1234	COP
1236	1244	STR	-	1240	1244	COP
1254	1263	STR	-	1261	1263	COP
1732	1738	STR	-	1732	1746	COP

ASPECTOS CURRICULARES PRECEDIENDO A COP:

En el primer texto analizado, la hipótesis no se cumple, pues en una de las ocasiones el código COP estuvo precedido de un código CUR. En el segundo caso, seis de las ocho apariciones de COP estuvieron precedidas por STR, y nunca por CUR, lo cual permitiría afirmar que en este texto se confirma la hipótesis de partida.

CONCLUSIÓN

Si bien la utilidad de los programas para el análisis de datos cualitativos de cara a la codificación, recuperación de fragmentos codificados y examen del contenido de las categorías es manifiesta, AQUAD permiten ir más allá de estas funciones añadiendo la posibilidad de comprobar hipótesis formuladas por el investigador. Con ello, se consigue una cierta objetividad y sistematización en el proceso, que podría ser desarrollado por distintos autores de modo que los resultados obtenidos para cada una de las posibles hipótesis serían coincidentes.

A pesar de todo, es preciso destacar que el tipo de hipótesis que permite formular y comprobar es limitado; las relaciones que pueden establecerse entre los elementos de significado se basan exclusivamente en la proximidad dentro del texto, la secuenciación, la inclusión o superposición de las áreas de texto codificadas. Otro tipo de relaciones no detectables por la disposición espacial de las unidades sino a partir del significado contenido en las mismas quedarían excluidas. Además, podría cuestionarse el que exista realmente una relación entre dos códigos que aparecen secuenciados, próximos o superpuestos. De ahí que se restrinja la formulación de hipótesis a los códigos que presentan cierta frecuencia, para que la reiterada aparición en el texto de determinados patrones de códigos apoye la existencia de una cierta relación entre los códigos implicados. Por otra parte, podrían existir dificultades para establecer cuándo una hipótesis ha sido confirmada. Como afirmábamos anteriormente, el contraste de hipótesis en este tipo de análisis no es un contraste en sentido estricto, y por tanto no tenemos criterios teóricos para rechazar o aceptar una hipótesis. Las decisiones pueden tomarse a veces sobre la base de una valoración del investigador acerca de la relevancia del número de apariciones de un código dentro de una estructura buscada frente al número total de apariciones de ese código. Los resultados de comprobar una hipótesis indican únicamente la evidencia encontrada para apoyar la plausibilidad de lo formulado.

En cualquier caso, consideramos que el programa AQUAD representa una interesante herramienta en manos del analista, de cara a la formulación de relaciones entre tópicos y a la búsqueda de evidencia que las confirme.

REFERENCIAS BIBLIOGRÁFICAS

- BOGDAN, R. C. y BIKLEN, S. K. (1982): *Qualitative research for education. An introduction to theory and methods*. Boston: Allyn and Bacon Inc.
- GARCÍA, E. et al. (1993): *Análisis de datos cualitativos en la investigación sobre la diferenciación educativa*. Ponencia presentada al VI Seminario de Métodos de Investigación, El Escorial.
- GIL FLORES, J. (1992): *Análisis de datos cualitativos. Aplicación al caso de los datos procedentes de grupos de discusión*. Sevilla: Tesis Doctoral inédita.
- HUBER, G. (1988): Análisis de datos cualitativos: la aportación del ordenador. En C. MARCELO (Coord.): *Avances en el Estudio del Pensamiento de los Profesores*. Sevilla: Servicio de Publicaciones de la Universidad de Sevilla. (pp. 77-85).
- HUBER, G. y MARCELO, C. (1990): Algo más que recuperar palabras y contar frecuencias: la ayuda del ordenador en el análisis de datos cualitativos. *Enseñanza*, (8), 69-84.
- LOFLAND, J. y LOFLAND, L. H. (1984): *Analyzing social settings*. Belmont, CA: Wodsworth Publishing Company.

NUDIST: UNA HERRAMIENTA INFORMÁTICA PARA EL ANÁLISIS DE DATOS CUALITATIVOS

Ana Corrales, Gregorio Rodríguez, Javier Gil y Eduardo García
Universidad de Sevilla

Desde la perspectiva de la metodología cualitativa la introducción del ordenador en el proceso de investigación ha supuesto un cambio radical y una nueva forma de pensar en el método cualitativo (Richards y Richards, 1991; García et al., 1993). Efectivamente, el uso del ordenador en el análisis de datos cualitativos de acuerdo con unas instrucciones (programas) que respondan a las necesidades de operaciones conceptuales que en cada caso se requieren, ofrece avances que van desde una mayor velocidad y capacidad de memoria, hasta las habilidades para localizar frecuencias de cadenas de caracteres en los textos pudiendo recuperarlas con rapidez mediante el uso de códigos, o llegar a la construcción de teorías, pasando por la comprobación de hipótesis cualitativas (Gil et al., 1993).

Actualmente existen numerosos programas informáticos que permiten realizar las operaciones requeridas en los procesos de análisis de datos cualitativos y que hasta la aparición de los mismos debían ejecutarse manualmente (Tesch, 1988, 1989, 1990, 1991; García et al., 1993). En el presente trabajo vamos a centrar nuestra atención en la descripción de las principales características y aportaciones que uno de ellos, concretamente el programa NUDIST (en su versión 2.3 para ordenadores compatibles), brinda al análisis de datos cualitativos.

EL SISTEMA NUDIST

Creado en la Universidad de La Trobe (Melbourne, Australia) NUDIST es un programa informático que nos permite manejar, organizar y realizar procesos de investigación en los que se manejan datos cualitativos de carácter textual, producto de entrevistas, observaciones, documentos históricos o literarios, notas de campo, noticias de periódicos, etc.

Siempre que se aborda el análisis de datos cualitativos es preciso acometer los siguientes procesos (Richards y Richards, 1992):

- Agrupar todos los documentos que se desean analizar en el proyecto;
- Indizar¹ segmentos del texto en varias categorías de indización;
- Buscar palabras y frases en los documentos;

¹ Los autores del programa insisten en la diferencia fundamental entre «indizar» y «categorizar», dando al primer término el sentido de una jerarquía de categorías, como podremos ver más adelante.

— Emplear la indización y búsqueda de texto como pilares básicos para encontrar párrafos en los textos y las ideas que en ellos se expresan;

— Elaborar notas y memos sobre las ideas principales y teorías a medida que se desarrolla el proyecto; y,

— Reorganizar y establecer la indización a medida que nuestros conocimientos y teoría aumentan en el proceso de investigación.

El programa NUDIST realiza estos procesos de forma eficaz y contiene, además, muchas otras características que no se encuentran en otros programas de análisis de datos cualitativos. De hecho, debido a su gran capacidad para indizar, buscar y teorizar —partes fundamentales del análisis cualitativo— este programa se ha llamado «NUDIST» que son la iniciales de «Non-Numerical Unstructured Data Indexing, Search and Theorizing» («Indización, búsqueda y elaboración de teorías de datos no numéricos y no estructurados»).

Para que el programa NUDIST trabaje eficazmente se crean un sistema de documentación y un sistema de indización, que son la base y el soporte de los distintos procesos que se ejecutan posteriormente.

En el **sistema de documentación** se puede almacenar, en primer lugar, cualquier cantidad de información de documentos textuales «on-line» (esto es, archivos en el propio disco del ordenador), con datos sobre los mismos que se deseen analizar a lo largo de la investigación. En segundo lugar se pueden almacenar documentos «off-line» (es decir, documentos que no están en el disco sino fuera del ordenador, como libros, notas, fotografías, etc.).

Ambos documentos (on-line y off-line) están divididos en unidades de texto como líneas, frases, diálogos, o cualquier otra opción se desee, siempre que se especifique claramente al programa NUDIST qué consideramos como unidad de texto. Asimismo, estos documentos y los datos acerca de los mismos, se pueden estudiar y visualizar de muy diversas formas. Podemos optar por su visualización a través de la pantalla del ordenador o salvarlos en un archivo para imprimirlo posteriormente, siempre con la cantidad de texto deseada.

Además, los documentos se pueden indizar bajo las categorías de indización que se deseen, pudiendo colocar encabezamientos (tan largos como se estime) para describir el contenido del documento, su origen ...etc., y subencabezamientos en el caso de los documentos on-line, también con cualquier longitud. Tanto el encabezamiento como el subencabezamiento aparecen con cualquier recuperación del documento del que procede.

El **sistema de indización**, por su parte, permite almacenar referencias a ideas, conceptos y hechos que surgen a lo largo de la investigación, y referencias a unidades de texto de los documentos on-line y off-line que muestren esas ideas, conceptos o hechos. El sistema de indización puede ser tan sencillo o complejo como el investigador lo desee o necesite, pudiendo contener cualquier cantidad de categorías de indización y, en cada categoría, las referencias de indización a las unidades de texto de los documentos de datos on-line y off-line que se deseen. Estas categorías, que pueden, además, tener subcategorías que a su vez pueden tener otras y así sucesivamente, se pueden crear durante los procesos de indización de los documentos, o bien mientras se avanza en la investigación, en un proceso continuo de mejora y adaptación del sistema de indización a las necesidades de la propia investigación.

Comprender lo que es un sistema de indización es fundamental para poder trabajar con NUDIST, aunque el programa lo ofrece como una posibilidad pudiendo realizarse los procesos de análisis tomando los códigos que se hayan elegido sin establecer ninguna relación entre los mismos, de ahí que profundicemos en algunos conceptos que son necesarios tener en cuenta a la hora de proceder al análisis de datos con este programa.

NUDIST permite organizar las categorías de indización en «árboles» de categorías, sub-categorías,

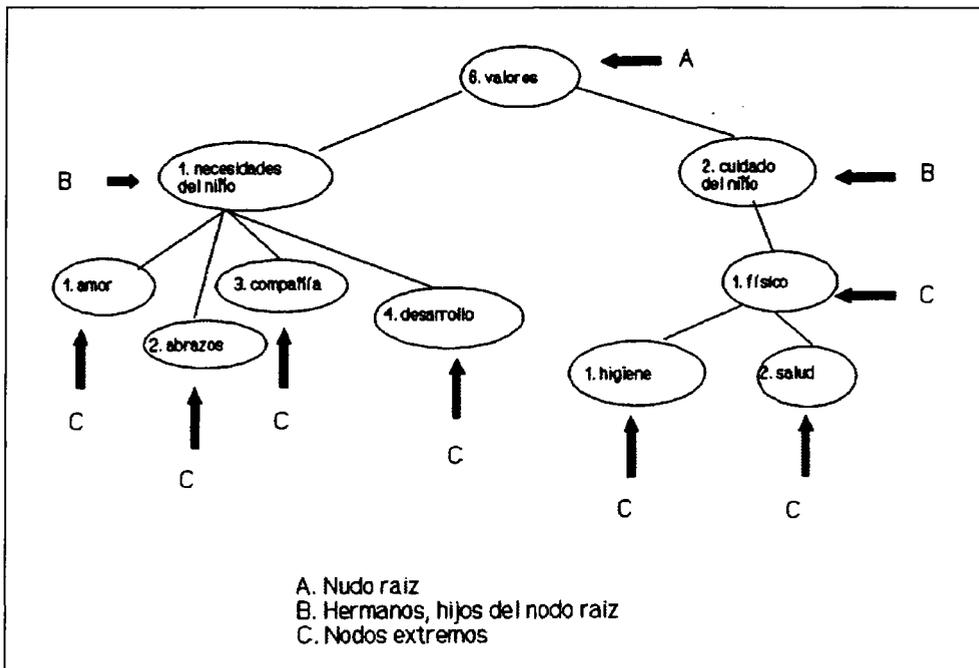


Figura 1. Ejemplo de árbol de indización.

sub-sub-categorías, etc. Los puntos que se encuentran al final y en la unión de las ramas son denominados **nodos**, no existiendo límite alguno en cuanto al número.

En la Figura 1 podemos contemplar una parte de un sistema de indización, en el que podemos observar cómo el nodo que se encuentra en el nivel superior recibe el nombre de **raíz**, no tiene a nadie por encima, pero sí puede tener «hijos» por debajo. Estos últimos que parten de los nodos raíces son conocidos como **hermanos**. Por último, los nodos que carecen de «hijos» son denominados **terminales o extremos**.

Cada nodo tiene un número, conocido como **número del nodo**, que van a permitir construir la **dirección del nodo**, a partir de la estructura jerárquica. Así, el nodo 1 2 1, se corresponde con el título higiene. En lugar de números, también cada nodo puede identificarse por el **título**, de esta forma al nodo 1 2 1 le corresponde el título *valores/cuidado del niño/higiene*.

La creación de las categorías de indización se puede llevar a cabo de tres formas diferentes:

- a) Tecleando las referencias de indización de las unidades de texto de un documento.
- b) Mediante la realización de una búsqueda automática de texto, localizando palabras, patrones de palabras o expresiones en los documentos on-line, empleando NUDIST para añadir las referencias resultantes a la base de datos de indización.
- c) Mediante el uso de una extensa gama de órdenes que permite crear nuevas categorías de indización a partir de las ya existentes.

La base de datos de indización puede ser de cualquier complejidad, desde una simple lista de códigos hasta estructuras en árbol de categorías y subcategorías de indización de una enorme complejidad. Esta estructuración permite organizar y manejar los conceptos indizados como sistemas teóricos, no simplemente como etiquetas.

Mediante el sistema de indización se puede:

- a. Asignar categorías de indización referidas a documentos completos, pudiendo almacenar información ilimitada sobre el documento.
- b. Almacenar comentarios en cualquier categoría de indización, grabar ideas sobre esa categoría o, por ejemplo, relaciones de esa categoría con otras.
- c. Recorrer el sistema de indización a través de la pantalla y guardar en un archivo cualquier operación realizada para imprimirlo posteriormente.
- d. Cambiar la indización de cualquier documento o cambiar libremente las categorías de indización cuando se desee.
- e. Recuperar cualquier parte de los documentos empleando órdenes que permiten cambiar las categorías de indización para expresar preguntas e hipótesis.
- f. Los resultados de la búsqueda de texto se pueden salvar en una categoría de indización. Esto significa que los resultados de búsqueda de texto, la indización manual y la construcción de nodos están completamente integrados en la base de datos de indización, y que pueden ser utilizados como base para estudios posteriores.

LOS PROCESOS QUE NUDIST NOS FACILITA

1. Navegar a través del sistema de indización

El sistema de indización es la vida y el alma de un proyecto de investigación que se sirva de NUDIST, de tal forma que el programa nos permite utilizar distintas herramientas para poder analizarlo, estudiarlo y, en su caso, modificarlo. En cualquier etapa del proceso de investigación en el que nos encontremos se puede hojear cualquier categoría o grupos de categorías en el sistema de indización, lo que se indiza y qué otras categorías se refieran a éstas.

2. Integrar la búsqueda de texto y la indización

La mayoría de los programas especializados en el análisis de datos cualitativos, o bien permiten la codificación y recuperación, o bien realizan búsqueda de texto. Sin embargo, el programa NUDIST combina la exploración de información de indización con búsquedas de expresiones que ocurren en el texto, así se puede buscar cualquier secuencia de caracteres en los documentos y recuperarla con la cantidad de texto deseada (la sección completa o incluso el documento entero en donde aparece).

Además, NUDIST permite guardar el resultado de una búsqueda de texto en un nodo del sistema de indización como indización de referencias de lo encontrado. Como cualquier otro nodo, se le puede asignar un nombre, incluir comentarios en él y contenido con información de indización sobre las unidades de texto de los documentos. Además, se puede hojear, alterar mediante la supresión o inclusión de algunas unidades de texto, mover o combinar en cualquier operación para construir nuevas categorías a partir de las ya existentes.

3. Comentarios en el sistema de indización

Lo importante al apoyar el desarrollo de teorías sobre los datos es la posibilidad de realizar comentarios sobre las categorías de indización, y éstos pueden proceder de dos fuentes principales: el investigador y el propio programa. Añadiendo, o borrando, en cualquier categoría de indización comentarios de textos el investigador dispone de una forma adecuada de almacenar notas de trabajo, memos, opiniones y pensamientos acerca de las categorías de indización. A su vez, NUDIST añade

automáticamente su propio comentario a las categorías de indización para grabar el momento de su creación o cualquier otro evento importante en su historia, proporcionando de esta forma una revisión continua del desarrollo de las ideas en el proceso de investigación.

4. Modificar el sistema de indización

Mientras la mayoría de los sistemas manuales y los programas de codificación-recuperación permiten, en cierta medida, recodificar los documentos, NUDIST está diseñado para controlar con una gran flexibilidad el sistema de indización y todas las indizaciones realizadas en el mismo. Efectivamente, el sistema de indización en NUDIST se puede hojear, explorar, reorganizar, cambiar y crecer a medida que nuevos conceptos e ideas sobre los datos emergen.

El objetivo es, pues, la máxima flexibilidad y control sobre la información. El investigador puede modificar el sistema de indización sin que tenga que recurrir de nuevo a los documentos previos; toda la información del sistema de indización puede recuperarse junta para ser recolocada en nuevos patrones de indización. Es más, determinados conceptos específicos se pueden agrupar automáticamente para poder explorar conceptos generales, incluso cuando estos últimos no han sido codificados explícitamente. Si la localización de un nodo (categoría de indización) en el sistema global no está en el sitio adecuado o deseado, NUDIST permite recolocarlo, con la información de indización, en el lugar donde proceda, y todos los nodos y subnodos siguientes pueden ser combinados, borrados o recolocados.

Así pues, esta flexibilidad permite que NUDIST se muestre como una útil herramienta de apoyo a las investigaciones que se inscriben en línea con la *teoría fundamentada* (Glaser y Strauss, 1967; Straus, 1987; Strauss y Corbin, 1990), método que se centra en la captura e interrogación de los significados emergentes de los datos.

5. Buscar y recuperar

Mediante la utilización combinada de una amplia variedad de operadores booleanos (Y, O y NO, etc.) y, a su vez, combinados con una serie de operadores no booleanos, se pueden establecer relaciones entre las unidades del texto y las estructuras de indización. NUDIST examina cuidadosamente los documentos recuperados para ayudar a manejar eficazmente la gran cantidad de información que se haya guardado en los documentos. Con cada recuperación de texto aparece el encabezamiento (en el que se ha grabado información sobre el documento), subencabezamientos referidos a cada pasaje recuperado (en los que se han podido anotar cuestiones como quién está hablando, las preguntas realizadas, etc., por ejemplo, en el caso de que se tratase de una entrevista), estadísticos numéricos sobre la recuperación y documentación que NUDIST ofrece sobre la historia de cada recuperación realizada.

6. Crear nuevas categorías de indización a partir las anteriores

NUDIST permite búsquedas indefinidas de ideas que surgen con posterioridad porque el resultado final de cualquier recuperación es un nuevo nodo en el sistema de indización, que puede ser explorado cuando se desee. Cada nuevo nodo que se construye está dotado de toda la información que acompaña a cualquier nodo: título, comentarios, fecha, facilidades de hojear, estadísticas, ...etc. De esta forma, al igual que en el modelo de la teoría fundamentada, el análisis que nos permite NUDIST se convierte en un proceso de ir explicitando las ideas emergentes, permitiendo tanto la construcción de teoría, como la comprobación de la misma.

7. Ficheros de comando para grandes cantidades de procesos

En lugar de realizar las operaciones a través de los menús que NUDIST facilita, éstas se pueden controlar y ejecutar a través de ficheros de comandos que el investigador construya. Por ejemplo, si se han de introducir numerosos archivos, o realizar búsquedas de texto complejas en varios archivos a la vez, o asignar códigos a determinadas unidades de texto, NUDIST permite hacerlo siguiendo las instrucciones escritas en un archivo de comandos en vez de tener que realizarlo a través de las selecciones del menú.

Esencialmente los archivos de comando pueden utilizarse para introducir documentos, crear nuevos árboles o sub-árboles de indización, buscar en los archivos on-line textos o patrones de texto y clasificar automáticamente lo hallado o ejecutar comandos para construir nodos y así crear nuevas categorías de indización a partir de la anteriores.

8. Integrar análisis de datos textuales y numéricos

Además de la lógica integración que podemos establecer entre, por ejemplo, los datos de identificación como sexo, edad, etc., que se suelen recoger en numerosas investigaciones, un aspecto importante de NUDIST es la posibilidad que nos ofrece de una serie de datos «estadísticos», como producto de las búsquedas realizadas, que posteriormente pueden ser utilizados para su análisis a través de procedimientos estadísticos.

9. Comprobación de hipótesis

Una de las posibilidades que nos ofrecen los ficheros de comandos es la comprobación de hipótesis de causalidad o de relaciones a lo largo de los textos objeto de investigación, entendiendo por hipótesis «las relaciones entre tópicos identificados en un texto a partir de las cuales es posible reconstruir un todo estructurado que nos aproxima a una comprensión global de los datos» (Gil et al. 1993: 1).

TRABAJANDO EN EQUIPO CON NUDIST

Aunque NUDIST 2.3 se sirve de menús y diálogos para que el investigador pueda controlar sus acciones, empleando sofisticadas formas para hojear, todas sus presentaciones se ofrecen por pantalla, simulando a un terminal de salida simple. Incluso empleando un Macintosh o Window 3.0 en un ordenador compatible, la ventana simple en la que NUDIST 2.3 funciona mantiene la misma estructura, de tal forma que el trabajo se asemeje con independencia del tipo de ordenador que se esté utilizando.

El programa NUDIST puede ejecutarse bien de forma individual o bien en equipo de trabajo. Tanto en versión para PC/Windows o Macintosh, se puede utilizar de forma individual, de tal forma que un único investigador puede controlar todo el proceso de investigación.

Pero si se utiliza una red local, o bien la versión para Unix o VAX/VMS, entonces existe la posibilidad de que varios investigadores compartan el mismo proyecto de investigación, aún cuando cada uno puede asumir distintas responsabilidades dentro del mismo.

BIBLIOGRAFÍA

GARCÍA, E. et al. (1993): *Análisis de datos cualitativos en la investigación sobre la diferenciación educativa*. Ponencia presentada al VI Seminario de Métodos de Investigación, El Escorial.

- GIL, J. et al. (1993): *Formulación y comprobación de hipótesis cualitativas con ayuda del programa AQUAD*. Comunicación presentada al VI Seminario de Métodos de Investigación, El Escorial.
- GLASER, B. G. y Strauss, A. L. (1967): *The Discovery of Grounded Theory*. Nueva York: Aldine Publishing Company.
- RICHARDS, L. y RICHARDS, T. J. (1991): The Transformation of Qualitative Method: Computational Paradigms and Research Processes. En N. Fielding y R. Lee (Eds.): *Using Computers in Qualitative Research*. Berkeley: Sage, (pp. 38-53).
- RICHARDS, T. et al. (1992a): *NUDIST 2.3. User Manual*. Eltham: Replee P/L.
- RICHARDS, T. et al. (1992b): *NUDIST 2.3. Reference Manual*. Eltham: Replee P/L.
- STRAUSS, A. L. (1987): *Qualitative Analysis for Social Scientist*. MA: Cambridge University Press.
- STRAUSS, A. L. y Corbin, J. (1990): *Basics of Qualitative Research. Grounded Theory Procedures and Techniques*. Newbury Park: Sage.
- TESCH, R. (1988): *The Impact of the Computer on Qualitative Data Analysis*. Comunicación presentada en la convención anual de la AERA, Nueva Orleans.
- TESCH, R. (1989): Basic Qualitative Analysis with QUALPRO. *International Journal of Qualitative Studies in Education*, 2 (4), 367-371.
- TESCH, R. (1990): *Qualitative Research: Analysis Types and Software Tools*. Hampshire: Falmer Press.
- TESCH, R. (1991): Software for Qualitative Researchers: Analysis Needs and Program Capabilities. En N. G. Fielding y R. M. Lee (Eds.): *Using Computers in Qualitative Research*. Londres: Sage Publications.

NUDIST VERSUS AQUAD: DOS POSIBILIDADES INFORMÁTICAS PARA EL ANÁLISIS DE DATOS CUALITATIVOS

Gregorio Rodríguez, Javier Gil, Ana Corrales y Eduardo García
Universidad de Sevilla

Desde que en 1966 apareciera el primer programa para el análisis de textos, el denominado *General Inquirer* (Stone et al., 1966), hasta la fecha muchos han sido los avances habidos en este campo, sobre todo en los últimos años de la década de los 80 y estos primeros 90.

Tesch (1988) presentaba ya un primer análisis comparativo de los programas de análisis de datos cualitativos disponibles en el mercado en ese momento, haciendo referencia a programas como THE ETHNOGRAPH, QUALPRO, QUALOG, TAP... Esta misma autora, en 1990 nos presenta un excelente trabajo en el que sintetiza éstos y otros programas, recogiendo sus principales funciones (Tesch, 1990). En esta segunda revisión puede observarse una rápida evolución en cuanto a la cantidad de programas, así como en la prestaciones que ofrecen y su disponibilidad. Tan sólo tres años después de la publicación del texto de Tesch, algunos programas han sufrido una gran transformación; éste es el caso de programas como NUDIST¹ (*Non-numerical Unstructured Data Indexing, Search and Theorizing* —Indización, búsqueda y elaboración de teorías de datos no numéricos y no estructurados) o AQUAD (*Analysis of Qualitative Data*). El primero se presenta en 1992 en su versión para ordenadores personales, en versiones para PCs y Macintoshes, y tiene ya anunciada su versión 3.0 para finales de este mismo año. El segundo, referido al caso concreto de nuestro país, en su versión 3.0 en castellano presenta mejoras sustanciales con respecto a las anteriores.

A lo largo de este trabajo vamos a centrar nuestra atención en estos dos programas (NUDIST y AQUAD) por ser considerados actualmente los dos programas que nos sirven como medio para la construcción y comprobación de teoría (Huber, 1990; Tesch, 1990; Richards, 1992).

Aunque AQUAD surge con posterioridad a NUDIST, en el caso de nuestro país es el programa que más se ha venido utilizando y ello por varios motivos. En primer lugar, por la difusión que se ha hecho del mismo: su creador —Günter Huber— ha impartido cursos en varias universidades españolas, exponiendo los fundamentos del programa e iniciando al manejo del mismo; ha participado en reuniones científicas celebradas en nuestro país (Huber, 1988) o ha publicado en revistas nacionales artículos sobre el tema (Huber y Marcelo, 1990). En segundo lugar por ser un programa que exige pocos requerimientos de hardware y está en castellano.

¹ Para una visión más amplia sobre NUDIST puede consultarse la comunicación de A. Corrales et al. (1993) «NUDIST»: *Una herramienta informática para el análisis de datos cualitativos* presentada en este Seminario.

NUDIST, por contra, hasta 1992 no ha dispuesto de una versión para ordenadores personales, lo cual ha limitado su expansión. Asimismo, en esta nueva versión los requerimientos de hardware son bastante elevados, representando ello un obstáculo para su utilización.

Independientemente de estas primeras diferencias, lo que realmente nos interesa en estos momentos son las posibilidades y limitaciones de cada uno de los programas de cara al análisis de datos cualitativos, y en ello centraremos nuestra atención, aunque con las limitaciones propias que la longitud fijada en este trabajo nos impone, para lo cual presentaremos un breve análisis comparativo de ambos programas.

1. ¿CÓMO TRABAJAN ESTOS PROGRAMAS?

En una primera aproximación podríamos decir que son bastante similares, por cuanto que a través de los mismos podemos llevar a cabo los principales procesos que todo análisis de datos cualitativos exige (García et al., 1993). NUDIST y AQUAD facilitan la tarea de análisis de datos cualitativos permitiendo codificar unidades de significado diferenciables en el texto, efectuar recuentos de tales unidades, examinar el contenido de las categorías resultantes o comprobar hipótesis. No obstante, desde el comienzo del trabajo con uno u otro programa existen ya diferencias apreciables.

a) Preparación del texto para el análisis

La base del análisis la constituye el texto con el que se trabaja, que para ambos programas debe presentarse en ficheros con formato ASCII, aunque con exigencias diferentes según veremos a continuación.

En el caso de AQUAD, es preciso numerar las líneas de texto, operación que el programa realiza automáticamente. Las líneas del texto han de estar formateadas a una longitud máxima de 48 caracteres (pulsaciones) por línea, de tal forma que quede en la pantalla espacio suficiente para agregar códigos a los márgenes, o bien sin limitación cuando se opta por introducir la localización de los códigos directamente en un fichero de códigos. El texto puede tener una longitud de 9999 líneas, alcanzando el límite de 64 KBytes de Turbo Prolog para strings (cadenas de signos). No obstante, pueden ser analizados múltiples textos a la vez, por lo que bastaría con segmentar en varios ficheros aquellos textos que rebasen el tamaño fijado.

NUDIST, por su parte, no tiene casi limitación alguna en la longitud de la unidad de texto, que puede llegar a ser de 32000 caracteres por párrafo. Si se opta por tomar como unidad de texto la línea habría que formatear el texto a 69 caracteres por línea; pero también puede optarse por tomar como unidad de texto la frase, el párrafo, la intervención (en un diálogo) o cualquier otra forma que se desee. A diferencia de AQUAD, NUDIST no permite realizar la codificación de forma directa sobre el propio texto en pantalla.

Otra diferencia por lo que se refiere al tipo de texto que permite analizar, es lo que NUDIST denomina ficheros «off-line», es decir, fichero donde se hace referencia a archivos, fotos, cintas o cualquier otro objeto que no se encuentra físicamente en el fichero del ordenador.

b) Codificación del texto

La codificación del texto implica asignar a cada unidad de significado diferenciable en el texto, un código que haga referencia al contenido de dicha unidad, y he aquí donde surge la primera gran diferencia entre ambos programas.

En NUDIST este proceso es denominado «sistema de indización», por cuanto que lo que se crean

son índices donde se ordenan jerárquicamente todas las categorías, y no se pone restricción alguna en cuanto al número de categorías del sistema o al número de categorías de una unidad de significado, pudiendo comenzar por unas pocas e ir construyendo el sistema poco a poco ya que pueden modificarse en cualquier momento.

Cuando codificamos un texto mediante el programa AQUAD, utilizamos códigos y metacódigos. Los códigos son asignados a determinados fragmentos de texto añadiendo junto a la línea donde comienza el fragmento codificado una etiqueta de identificación seguida del número de línea donde finaliza el fragmento. Por ejemplo, si quisiéramos codificar un fragmento de texto comprendido entre las líneas 15 y 19, en el que se habla acerca de la evaluación, utilizando para ello el código EVA, escribiríamos al final de la línea 15 la expresión VA-19. Un mismo fragmento puede ser codificado con varios códigos, puede estar incluido o superpuesto parcialmente con un fragmento correspondiente a otro código. Los fragmentos correspondientes a una serie de códigos vinculados por algún criterio, pueden ser codificados automáticamente por un metacódigo.

Tanto NUDIST como AQUAD permiten efectuar un recuento de las apariciones de códigos en los textos analizados, haciendo posible el análisis de los datos mediante procedimientos tales como el análisis secuencial, los modelos lineales logarítmicos, el análisis factorial de correspondencias y el escalamiento no métrico, el análisis de correlación canónica, los modelos logit, etc. (Ver Ponencia sobre «Análisis de datos cualitativos en la investigación sobre la diferenciación educativa»).

c) Examinar el contenido de las categorías y los sistemas de codificación

En NUDIST las recuperaciones de texto son ilimitadas. No hay restricciones en el número de veces que un pasaje de texto puede ser indizado, y virtualmente ninguna restricción en las preguntas que puede hacerse sobre la indización de líneas de un texto, o la indización en archivos o la información sobre el propio archivo. Se pueden utilizar operadores booleanos en cualquier combinación de unos con otros, y combinados con operadores no booleanos que establecen relaciones de los pasajes en el texto y estructuración de indización.

Además, NUDIST permite examinar los propios documentos y el sistema de indización de forma integrada, permitiendo de esta forma al usuario el control para reorganizar y reformar todos los aspectos del sistema. En NUDIST, el sistema de indización puede ser ojeado, explorado, reorganizado y cambiado, creciendo a medida que nuevos conceptos emergen de los datos.

El principal objetivo de AQUAD no es la organización de sistemas de categorías jerarquizados, por lo que su flexibilidad para renombrar códigos es mucho menor. En cambio, sí contribuye eficazmente a la revisión del contenido de las categorías, permitiendo la recuperación de todos los fragmentos etiquetados con determinados códigos.

d) La comprobación de hipótesis

Considerando la comprobación de hipótesis con las limitaciones que han de entenderse desde el análisis de datos cualitativos por ordenador —búsqueda de evidencia que apoye una afirmación—, tanto AQUAD² como NUDIST permiten esta comprobación, aunque con diferencias.

AQUAD ofrece un repertorio de hipótesis sobre vínculos existentes entre categorías, basadas todas ellas en relaciones de secuenciación, inclusión, superposición o coaparición dentro de una distancia

² En el caso concreto del programa AQUAD puede consultarse la comunicación de J. Gil et al. (1993): *Formulación y comprobación de hipótesis cualitativas con ayuda del programa AQUAD*, presentada en este Seminario.

CUADRO 1
 CARACTERÍSTICAS DE LOS PROGRAMAS NUDIST Y AQUAD

	NUDIST 2.3	AQUAD 3.0
Lenguaje de programación	• LISP	• Turbo-Prolog
Versiones	• VAX / PC-386 /Mac	• PC
Hardware necesario	• PC-386DX o SX • 4 Mb (RAM); 2 Mb (HD)	• PC XT • 640 K (RAM)
Entrada	• Fichero ASCII	• Fichero ASCII
Límites de los ficheros	• Ilimitado	• 64 K por fichero
Comandos	• Todos los comandos están disponibles en pantalla • Ficheros de comandos	• Comandos disponibles en pantalla
Principales procedimientos	• Indizar, codificar, recuperar, buscar, contar, numerar líneas, comprobación de hipótesis.	• Numerar líneas, codificar, recuperar, contar, comprobar hipótesis
Códigos	• Sistema de indización en árbol • Nodos ilimitados	• Sistema de códigos y metacódigos • Posibilidad de codificación múltiple
Procedimiento de codificación	• Directamente • Por búsqueda de texto • A través de comandos	• Directamente • A través de comandos
Modificación de la codificación	• Borrar, modificar, cambiar, sustituir, añadir...	• Añadir metacódigos
Salida de datos	• Pantalla • Fichero ASCII	• Pantalla • Impresora • Fichero ASCII
Ventajas fundamentales	• Flexibilidad • Inicio de codificación automática	• Formulación de hipótesis propias
Inconvenientes más graves	• Manual y programa en Inglés • Requerimiento de Hardware	• Excluye comprobación de relaciones no basadas en coaparición de códigos
Manual	• Guía para comenzar • Manual del usuario • Manual de Referencia	• Manual
Precio	±24.000 Ptas	±20.000 Ptas.

máxima de los fragmentos correspondientes a dos o más códigos. Permite además al analista formular sus propias hipótesis en lenguaje turbo-prolog, para buscar evidencia de la aparición de diferentes constelaciones de códigos que se repiten a lo largo del texto.

En NUDIST la comprobación de hipótesis puede llevarse a cabo de forma fácil a través de los ficheros de comandos, permitiendo contrastar la proximidad de determinadas categorías o estableciendo ciertas relaciones causales o de otro tipo.

e) La rapidez en el trabajo

A fin de evitar un trabajo monótono en el que sería preciso repetir muchas operaciones, NUDIST prevee la ejecución de tales tareas a partir de la realización de ficheros de comandos, a través de los cuales se puede ejecutar acciones como introducir los documentos que se van a analizar, ejecutar procesos de búsqueda guardando los resultados en un nodo, crear nodos y darles títulos, crear árboles enteros, indizar un documento, ejecutar operaciones de creación de nodos más complejas que las realizadas de forma interactiva. En definitiva, a través de este sistema se pueden realizar las operaciones más fundamentales e importantes, sin necesidad de tener que estar interactuando continuamente con el ordenador.

En el caso de AQUAD, es necesario ir seleccionando las operaciones que deseamos realizar a partir de un menú visualizado en la pantalla.

2. A MODO DE CONCLUSIÓN

En forma de síntesis, en el cuadro 1 ofrecemos una información básica sobre las principales diferencias existentes entre los dos programas analizados, y aun cuando consideramos que es el propio usuario quien debe tomar las decisiones en función de las características y objetivos de su investigación, podemos decir que NUDIST es un programa más flexible que AQUAD, que nos permite trabajar con mayor cantidad y diversidad de formatos de texto y de forma más rápida. Pero también adolece de unas limitaciones como son el estar íntegramente en inglés y con unas presentaciones en pantalla que no aprovecha el entorno gráfico de Windows, limitación esta última que parece que se subsanará en la nueva versión 3.0 de la que se dispondrá a finales de 1993, lo cual le hará situarse entre los programas que trabajan como una aplicación bajo Windows. Y en cuanto al requerimiento de hardware, el que precisa ya se viene convirtiendo poco a poco en estándar.

Como hemos podido comprobar a lo largo del presente trabajo, las operaciones que se suelen realizar en los procesos de análisis de datos cualitativos quedan suficientemente facilitadas por cualquier de los dos programas informáticos analizados. No obstante, si el investigador está interesado en llevar a cabo una asignación de códigos jerarquizados, con lo que ello conlleva de supuestos implícitos, no cabe duda que su elección ha de dirigirse hacia el programa NUDIST.

REFERENCIAS BIBLIOGRÁFICAS

- CORRALES, A. et al. (1993): «*NUDIST*»: *Una herramienta informática para el análisis de datos cualitativos*. Comunicación presentada al VI Seminario de Métodos de Investigación, El Escorial.
- GARCÍA, E. et al. (1993): *Análisis de datos cualitativos en la investigación sobre la diferenciación educativa*. Ponencia presentada al VI Seminario de Métodos de Investigación, El Escorial.
- HUBER, G. (1988): *Análisis de datos cualitativos: la aportación del ordenador*. En C. MARCELO (Coord.): *Avances en el Estudio del Pensamiento de los Profesores*. Sevilla: Servicio de Publicaciones de la Universidad de Sevilla, (pp. 77-85).

- HUBER, G. y MARCELO, C. (1990): Algo más que recuperar palabras y contar frecuencias: la ayuda del ordenador en el análisis de datos cualitativos. *Enseñanza*, (8), 69-84.
- RICHARDS, L. y RICHARDS, T. J. (1991): The Transformation of Qualitative Method: Computational Paradigms and Research Processes. En N. Fielding y R. L. (Eds.): *Using Computers in Qualitative Research*. Berkeley: Sage, (pp. 38-53).
- RICHARDS, T. et al. (1992). *NUDIST 2.3. User Manual*. Eltham: Replee P/L.
- RICHARDS, T. et al. (1992): *NUDIST 2.3. Reference Manual*. Eltham: Replee P/L.
- STONE, P. J., DUNPHY, D. C., SMITH, M. S. y ORGILVIE, D. M. (1966): *The GENERAL INQUIRER: A computer approach to content analysis*. Cambridge, MC: The M.I.T. Press.
- TESCH, R. (1988): *The Impact of the Computer on Qualitative Data Analysis*. Comunicación presentada en la convención anual de la AERA, Nueva Orleans.
- TESCH, R. (1989): Basic Qualitative Analysis with QUALPRO. *International Journal of Qualitative Studies in Education*, 2 (4), 367-371.
- TESCH, R. (1990): *Qualitative Research: Analysis Types and Software Tools*. Hampshire: Falmer Press.
- TESCH, R. (1991): Software for Qualitative Researchers: Analysis Needs and Program Capabilities. En N. Fielding y R. L. (Eds.): *Using Computers in Qualitative Research*. Berkeley: Sage, (pp. 16-37).

CATEGORIZACIÓN Y COMPLEJIDAD EN LA INVESTIGACIÓN CUALITATIVA

Xavier Gil Quesada
Universidad de Barcelona

0. INTRODUCCIÓN

El **análisis cualitativo** pretende, entre otras cosas, dar sentido y significado a datos que admiten un formato narrativo. Es importante tener presente que **significado** es un término estructural ya que hace referencia a las propiedades emergentes del sistema sobre el que se centra la atención del investigador. Es decir, del **todo** se extrae la significación que no se hallaría en el estudio aislado de cada una de las partes.

En el análisis cualitativo los datos se segmentan en unidades significativas; y los segmentos son categorizados según un sistema organizativo que surge de los propios datos. Con ello, el resultado del análisis es una síntesis de alto nivel. Es absolutamente necesario plantearse el papel de la complejidad como estrategia para la descripción e interpretación de la realidad.

Dos tipos de análisis se complementan interactuando: el estructural y el interpretativo. Mientras la estructuración pretende el hallazgo y establecimiento de relaciones entre las partes, la dimensión interpretativa ha de segmentar (descontextualización) para categorizar en la búsqueda de una nueva complejidad. Esta comunicación se centra en la estructuración (que se propone compleja) y en la categorización que asume la lógica difusa como vía adecuada para el tratamiento de la complejidad.

1. SOBRE EL CONCEPTO DE CATEGORIZACIÓN

Según Tesch (1990), categorización puede considerarse un término sinónimo de clasificación. Partiendo de esta identificación, lo primero que hemos de constatar es que clasificar es una forma de conocer. Para poder clasificar, la Teoría Clásica de Conjuntos exige el establecimiento de una relación de equivalencia. Cada clase está formada por todos los elementos relacionados (equivalentes), y todas las clases (conjunto cociente) forman una partición del conjunto en el que se ha definido la relación: son disjuntas y lo recubren. Cada clase tiene un representante canónico o una etiqueta que la representa. Es éste un modelo simple y rígido de categorización que parte de una equivalencia y finaliza en su codificación.

En este modelo juega un papel fundamental y restrictivo la propiedad transitiva. Una parte considerable de la realidad educativa quedaría excluida de los modelos de categorización con esta exigencia: no siempre podemos sostener que, cuando un alumno (o un contenido, un ítem o un tratamiento) es similar a otro y éste lo es a un tercero, el primero y el último también lo son.

Ya en 1901, Poincaré (citado por Valverde 1985) afirmaba que la indistinguibilidad física no es transitiva. Para el tratamiento de realidades complejas como la educativa, es necesario introducir modelos de categorización más amplios y flexibles. El uso de proximidades y similitudes (que admiten una cierta gradación en la transitividad) posibilitan agrupaciones más isomorfas a los sistemas que son objeto de nuestro estudio. Estas categorías también recubren todo el dominio como en los modelos clásicos, pero su riqueza estriba en permitir un cierto solapamiento de las «clases» y unos límites difusos que admiten la gradación en la pertenencia (Tesch, 1990).

Con unos modelos más flexibles como los descritos anteriormente, la codificación cobra una mayor riqueza en su función de simbolización y etiquetaje estructurado de las agrupaciones o categorías. Una entidad puede ser etiquetada con dos códigos distintos (solapamiento) aunque su grado de pertenencia a las categorías que cada uno de ellos representa pueda ser diverso. En este sentido podría ser metodológicamente clarificador la utilización de **árboles ponderados** (en la terminología de Bliss (1983) podrían denominarse fuzzy system networks).

2. ESTRUCTURACIÓN: HACIA UNA COMPLEJIDAD ORGANIZADA

El concepto de **complejidad** está emergiendo como fundamental en las ciencias humanas, tan fundamental tal vez como lo es el concepto de energía en las ciencias naturales. Su foco referente se halla en la Teoría de Sistemas (Bertalanffy, 1968). Formalmente entendemos por sistema un conjunto formado por objetos y por relaciones entre estos objetos. Estas relaciones pueden ser diversas y tienen el rol central en la definición de la estructura.

Existen muchas maneras de caracterizar la complejidad; pero en general el índice de complejidad está asociado tanto con el número de partes reconocidas como con el nivel y el tipo de sus interrelaciones. Si la lógica utilizada es difusa, el tipo de relaciones que aceptamos para las categorizaciones es de mayor complejidad que si nos restringimos a las clásicas. Por otra parte, se ha de tener en cuenta que la complejidad tiene una cierta connotación subjetiva, desde el momento en que está relacionada con la capacidad para entender o enfrentarse a un objeto.

Remitiéndonos a Klir (1985) proponemos la distinción metodológica entre simplicidad y complejidad, confrontada con el nivel de organización. En este contraste, la **simplicidad organizada** se caracteriza por un número de variables muy pequeño y un alto grado de determinismo. Simboliza en buena medida la historia de la ciencia moderna, en cuanto sus modelos hacen referencia a sistemas muy simples (dos o tres objetos). Se ha dicho, y con razón, que Newton es el gran simplificador. Este paradigma ha posibilitado grandes logros en el estudio de fenómenos físicos, pero ha presentado menor eficacia en ciencias como la biología y en medicina; y su incidencia en las ciencias sociales y humanas ha sido a penas perceptible.

Por otra parte, la **complejidad desorganizada** se caracteriza por un número de variables muy grande actuando de una manera altamente aleatoria. Boltzmann y Gibbs, a partir de los estudios de probabilidad anteriores, desarrollaron modelos estadísticos potentes para atacar este tipo de problemas, con lo que se consiguen grandes avances en mecánica estadística, termodinámica, genética, ingeniería, economía, etc. También dan un fuerte impulso a la investigación en las ciencias humanas y sociales.

Mientras los métodos analíticos, desarrollados por la simplicidad organizada, se hacen impracticables cuando aumenta el número de variables, la relevancia y precisión de los estadísticos mejora al aumentar este número. Estos dos tipos de métodos son, por tanto, muy complementarios. Cubren los dos extremos del espectro de la Complejidad, pero dejan al descubierto casi toda su zona central.

Esta región metodológica intermedia correspondería a lo que ya se ha denominado **complejidad organizada**. Su característica fundamental no es el número elevado de variables, sino el que se

presentan con un elevado nivel de organización. Por tanto no son susceptibles de ninguna de las dos simplificaciones clásicas de la ciencia.

De todos modos, no todos los grados de complejidad son abordables; si un problema presenta un grado de complejidad excesivo, habrá que reducirlo hasta que sea aceptable a la mente del que ha de abordarlo. En general, una simplificación será adecuada cuando minimice la pérdida de información significativa.

Hay motivos para pensar que el único modo de trabajar con sistemas que poseen complejidad organizada es permitir un cierto grado de **imprecisión** en la descripción de las agrupaciones de datos. Como estableció Zadeh (1973) parece necesario aceptar un principio de incompatibilidad: «... de manera un tanto informal, este principio dice que, a medida que la Complejidad de un Sistema aumenta, nuestra capacidad para hacer afirmaciones precisas y relevantes respecto a su comportamiento disminuye hasta un umbral más allá del cual precisión y relevancia son casi mutuamente incompatibles». Continúa Zadeh: «... los elementos clave en el pensamiento humano no son números, sino etiquetas o conjuntos difusos, es decir, clases de objetos en los que el paso de la pertenencia a la no-pertenencia es más gradual que abrupta». Esta lógica juega un papel básico en aquello que podría considerarse una de las más importantes facetas del pensamiento humano: la **capacidad para resumir información**, extrayendo de las colecciones de masas de datos que impactan sobre el cerebro aquellas y sólo aquellas sub-colecciones que son relevantes para la realización de la tarea que se ha de hacer.

CONCLUSIÓN

Categorizar es un proceso que trata de romper el corsé del rigor matemático en busca de una simplificación de la realidad que respete un cierto grado de complejidad estructurada. La descontextualización es un primer paso del camino que lleva, por medio de la segmentación/categorización, desde una complejidad excesiva a un modelo con una complejidad manejable. Para llevarlo a buen término, tal vez, como afirma Zadeh, necesitemos romper con «el fetichismo de la precisión, del rigor y del formalismo matemático, y utilizar una estructura metodológica que sea tolerante con la imprecisión y con la verdad parcial».

BIBLIOGRAFÍA

- BERTALANFFY, L. von (1968): *General System Theory*. G. Braziller. New York.
- BLISS, J. (et al.) (1983): *Qualitative Data Analysis for Educational Research*. Croom Helm. London, Camberra.
- KLIR, G. (1985): *Architecture of systems problem solving*. Plenum Press. New York.
- TESCH, R. (1990): *Qualitative Research: Analysis types and software tools*. The Falmer Press. New York.
- ZADEH, L. A. (1973): *Outline of a new approach to the analysis of the complex systems and decision processes*. I.TSM Cybern.

PROBLEMAS TÉCNICOS ENUNCIADOS EN EL ÁMBITO DE LA INVESTIGACIÓN ETNOGRÁFICA SOBRE LA DIFERENCIACIÓN EDUCATIVA

Fernando Sabirón Sierra
Universidad de Zaragoza

Dada una serie de informaciones y datos cualitativos obtenidos en un proceso de investigación etnográfica, en este caso sobre la diferenciación educativa, ¿cuáles son, habida cuenta de la evolución reciente de este tipo de investigación, los *problemas técnicos* no resueltos en la recopilación, tratamiento, análisis, utilización y distribución de este tipo de informaciones y datos?

Sabido es —sin necesariamente apelar por ello al tan tergiversado como difundido concepto khuniano de paradigma— que, *grosso modo*, cabe diferenciar tres fases sucesivas en la relación entre los modelos y las comunidades científicas que utilizan prioritariamente un tipo de investigación neopositivista o etnográfica¹. En una primera fase de oposición y enfrentamiento se utiliza no tanto un tipo u otro de modelos de investigación, sino que se desprecia y resaltan las limitaciones del tipo por el que no se ha optado, antes que desarrollar el propio. Los «problemas» predominantes en esta fase son estilísticos: dependen de lo accidental en la manera de aplicar, argumentar y defender la elección realizada por el investigador, con cierta independencia de lo sustancial de la investigación (rigor, adecuación, etc.). Dispondremos así, en esta primera fase, de un conocimiento científico reducido a entelequias y presentado, eso sí, como saber cifrado o novelado, según el tipo de investigación, pero siempre desatinado. La segunda fase en la relación que nos ocupa es mercantilista. En cada tipo de investigación aplicado a líneas particulares se acota un ámbito determinado de influencia sin interferir en el vecino. Es la fase de producción medida en número de publicaciones que aboca rápidamente en una alta inflación entre el valor y rigor real de las investigaciones y el atribuido en el «ranking» académico-universitario al uso. Recuérdese la década de los ochenta: tan patente fue la beligerancia pasionada que caracterizaba a los defensores de la investigación etnográfica, como las críticas de los «objetivos detractores» de la misma; para, en la actualidad, encontramos en un momento de sosiego en el que incluso no es excepcional el «transfugismo» de investigadores —obviamente y en puridad, circulación, transferencia entre corrientes, etc., califíquese como se prefiera— de un tipo a otro de modelos de investigación. Es el momento idóneo para iniciar, quizá, una tercera fase caracterizada por un desarrollo independiente de cada tipo de investigación. Si efectivamente se pretende el desarrollo y

1 Los términos utilizados en la denominación de un tipo u otro de investigación varían hasta límites insospechados de vaguedad y ambigüedad. Sin entrar en discusión, dado que no es el tema central de esta comunicación, se opta por los dos indicados.

progreso científico en nuestro ámbito de conocimiento, con la necesaria combinación —en su caso— de distintos tipos de investigación sobre un mismo objeto, esta tercera fase es inaplazable en la investigación etnográfica cuando nos encontramos, en la actualidad, en un punto en el que o bien se repiten ventajas e inconvenientes sin aportar solución ni alternativa alguna; o bien se opta por una posición erróneamente ecléctica en la que la investigación etnográfica es el adorno a la pedagogía experimental clásica. Las razones del «impasse» son, cómo no, múltiples y dispares —creencias, intereses, ideología, disponibilidad de medios, etc.— pero nos importan aquí, únicamente, las metodológicas. En la metodología cualitativa preferente en investigación etnográfica no se han resuelto todavía una serie de problemas técnicos superados ya en una metodología cuantitativa, con una tradición más fuerte en la investigación social.

La intención de esta comunicación es, sencillamente, enunciar algunos de los problemas principales —que no las soluciones— de la metodología cualitativa en un intento por colaborar en la clarificación del panorama. El foro de este «VI Seminario de Modelos de Investigación Educativa» no puede ser más apropiado.

I. REFERENTES EPISTEMOLÓGICOS Y TEÓRICOS EN INVESTIGACIÓN ETNOGRÁFICA SOBRE LA DIFERENCIACIÓN EDUCATIVA

La acotación de la investigación educativa queda delimitada por cinco referentes comunes, aunque de naturaleza distinta: de una parte, los referentes epistemológicos, teóricos y técnicos; de otra, los educativos y escolares. Estos cinco referentes se articulan, a su vez, en torno a dos ejes: homogeneizador y diferenciador (Berger, 1993). En una posible tesis de partida, ajustada a los problemas técnicos que se indicaban en párrafos anteriores, cabe establecer: (1) la asociación tanto de la metodología cualitativa, como de la investigación etnográfica en su totalidad, al eje diferenciador y (2) la obligación para el investigador en el ámbito de la diferenciación educativa, de ser consecuente con las exigencias de tal opción diversificadora. En caso contrario, y en el mejor de los supuestos, cuando menos la credibilidad de la investigación se resentirá.

El desglose de los referentes queda establecido de la siguiente manera esquemática:

(a) *Referentes epistemológicos en la investigación etnográfica de la diferenciación educativa*²: En primer término, la combinación de un deliberado enfrentamiento entre «paradigmas» (Séguin y Chanlat, 1987 con la posibilidad de «transferencia» entre «esquemas de interpretación» (Berthelot, 1987 respecto a la generación de conocimiento que permite mantener las connotaciones ideológicas del paradigma de referencia (Popkewitz, 1988, Guba y Lincoln, 1990 —inherentes, por otra parte, a todo paradigma (Masterman, 1975; Barnes, 1987— sin por ello excluir el conocimiento generado por otros modelos de investigación. De manera que, y en segundo término, *la zona de generación de conocimiento propia de la investigación etnográfica queda acotada por los modelos hermenéuticos, fenomenológicos y dialécticos*; frente a la preferencia por los modelos causales, funcionales y estructurales propios de la investigación neopositivista.

(b) *Referentes teóricos en la investigación etnográfica de la diferenciación educativa*: La «diferenciación educativa» considerada como objeto de investigación ha de combinar la diversificación de los «partenaires» educativos —anteponiendo, siempre y sin tópicos, el interés de la persona— con la mejora de la intervención escolar (*vid. infra.*). Téngase en cuenta que, aunque en ocasiones nos pese, la «diferenciación educativa» pasa, hoy por hoy y necesariamente, por una acción escolar. Será pues un análisis de las condiciones escolares las que conformen la interpretación e investigación de una

2 Para una exposición más detenida me remito a un documento anterior sobre el mismo tema (Sabirón, 1992a).

hipotética diferenciación educativa e integración social³. Los presupuestos de la diferenciación educativa aplicados y considerados desde teorías organizativas de los centros docentes abocan así, por necesidad, en intereses interpretativos, críticos y de mejora, frente a la eficacia y optimización de los presupuestos homogeneizadores⁴. Análogamente, los referentes teóricos estarán acotados por los modelos simbólicos (Mead, Rose, Blumer, Goffman, Schutz, ...) ⁵ y los socio-críticos (entre los más destacables, Weick, 1983).

II. FORMULACIÓN DE LOS PROBLEMAS A RESOLVER EN EL ÁMBITO DE LA INVESTIGACIÓN ETNOGRÁFICA SOBRE LA DIFERENCIACIÓN EDUCATIVA

Esbozados los referentes y desde una interpretación, sin duda sesgada, del estado actual de la investigación etnográfica en su conjunto, los problemas quedan enunciados como sigue:

(Epistemológicos)

1. La naturaleza intersubjetiva de la información interpretada en un proceso de investigación, el conflicto de intereses entre las partes, las expectativas de los participantes, el control por parte de los «protagonistas», la intención de emancipación, ... que exigen los «esquemas» hermeneúticos, fenomenológicos y dialécticos pueden acarrear situaciones comprometidas y de compromiso del investigador durante el proceso y ante los resultados de la investigación: ¿Qué posición deontológica, ética y legítima, cabe adoptar?

(Teóricos)

2. La adhesión a las teorías interpretativas y críticas proponen, entre otros presupuestos, la necesidad de partir de las prácticas para elaborar las teorías, la necesidad de contextualizar las acciones en y a cada situación, ... (2.1) ¿Es posible superar, en investigación etnográfica sobre la diferenciación educativa, la contradicción que supone la transposición literal de análisis teóricos generalizados a realidades escolares con un historia, una tradición, una configuración, unas normas, una «cultura educativa» sustancialmente diferentes y singulares? (2.2) ¿Es posible superar la transposición lineal de las teorías a las prácticas? En el caso de que estas dos reducciones se superaran: (2.3) ¿no se reforzarán las desigualdades sociales con intervenciones educativas estrictamente ajustadas a las condiciones del contexto escolar?; (2.4) ¿cómo evitar el riesgo de servir de legitimación «científica» a intenciones eficientistas de la política educativa? El dilema queda por resolver pues, pese a estas limitaciones, (2.5) cualquier otro tipo de investigación unida a una metodología prioritariamente cuantitativa, ya demostrado tanto la incapacidad manifiesta para llevar a cabo procesos de mejora efectiva para cada caso, como el recurso a la generalización para legitimar la homogeneización de la población y segregar —siempre «científicamente», por supuesto— al distinto.

3 Es ilustrativo en este sentido el análisis a la «confubalización de la ciencia y la ideología» sobre la escolarización del niño deficiente mental realizado por Molina (1992).

4 Cuestión ésta desvelada tajantemente en las revisiones a la «cultura escolar» que tuve el placer de analizar (Sabirón, 1992b).

5 Para una excelente síntesis consúltese a Morgan (1990).

(Técnico-Metodológicos)

3. La metodología cualitativa y el «estudio de caso», en sus distintas versiones, como estrategia de investigación no está exenta de una particular problemática técnica. (3.1) En primer término, es necesario combinar el diseño flexible, sometido a constantes reformulaciones, con el rigor en la recogida de la información. (3.2) En segundo término, siguen sin resolverse los problemas metodológicos de «credibilidad», «transferibilidad», «estabilidad» y «fiabilidad» ya señalados por Guba (1985). (3.3) Como siguen sin solución técnica la vulnerabilidad de las percepciones, el exceso de datos, la dificultad en la definición de roles, el proceso dialógico-dialéctico (Sabirón y Lozano, 1990). (3.4) En tercer lugar, los procesos de contraste entre instrumentos, determinantes para este tipo de investigación, siguen realizándose de manera un tanto burda. La cada vez más difundida utilización de «paquetes informáticos» para el tratamiento de la información cualitativa demuestra, como mínimo, cierta contradicción pues no sólo se corre el riesgo de reducir, *v. gr.*, la categorización a recuento, sino que se anula la profundidad de la información cualitativa obtenida por el posterior ajuste generalizador en el tratamiento con la lógica formal del ordenador. (3.5) En el contraste entre distintas perspectivas ciertamente se ha avanzado (Lozano, 1993, pero en el caso particular de investigaciones sobre diferenciación educativa, ¿estamos dispuestos a atribuir a los protagonistas la plena autonomía y responsabilidad en la posterior y definitiva toma de decisiones?

(Técnico-Educativos)

4. La «diferenciación educativa» demanda una obvia diversificación entre las partes. Si a la ya clásica complejidad organizativa de toda investigación o acción educativa se le añade el complejo entramado de instituciones y estamentos que intervienen en una acción diferenciadora y que, en consecuencia, será necesario abarcar en una investigación etnográfica, ¿siguen siendo viables los modelos holísticos? o será acaso necesario aplicar con rigor y desarrollar metodológicamente los, por ejemplo, «diagramas contextuales» (Huberman y Miles, 1991 más próximos a una intención comprensiva de sucesivas «aproximaciones a».

(Técnico-Escolares)

5. La finalidad implícita a todo proceso de investigación etnográfica —máxime en su aplicación a la «diversificación educativa»— estriba en la mejora de los procesos. En esta cuestión sí que existe un acuerdo generalizado, pero que se rompe ante la necesidad, en un caso dado, de determinar los criterios de tal mejora. Dicho de otro modo, cuando un proceso de investigación ha de abocar en un paralelo proceso comparativo de mejora, ¿cómo mantener la autonomía del individuo, grupo social o institución en la decisión del sentido, la pertinencia, la viabilidad, la oportunidad, la priorización, la instrumentación, etc., de tal mejora, con la intervención, pretendida o no, del investigador?. Si este dilema ya era lo suficientemente grave en ámbitos menos comprometidos como pudiera ser el de la evaluación de centros docentes en el que resultaba necesario el contraste entre el funcionamiento interno del centro, la previsión de los efectos en el entorno inmediato y la responsabilidad del centro frente a la justicia social (Derouet, 1992; Sabirón, 1993, ¿cómo respetar estas mismas condiciones en «grupos de discusión» (Morgan, 1988 relativos a la «diferenciación educativa»?

Sirvan estos problemas enunciados únicamente como pretexto para la discusión pues, en ningún caso, se pretende aportar solución decisiva alguna; sino, simplemente, seguir avanzando en el tratamiento de las informaciones y datos cualitativos y en el, a mi entender, inseparable adelanto de la investigación etnográfica.

REFERENCIAS BIBLIOGRÁFICAS

- BARNES, B. (1987): *T.S. Kuhn y las ciencias sociales*. México. FCE.
- BERGER, G. (1993): *Evaluación vs. control*. Seminario impartido en Zaragoza. Dpto. de Ciencias de la Educación. Universidad de Zaragoza.
- BERTHELOT, J. M. (1991): *L'intelligence du social*. Paris. PUF.
- DEROUET, J. L. (1992): *Ecole et Justice. De l'égalité des chances aux compromis locaux?* Paris. Métailié.
- GUBA, E. G. (1985): Criterios de credibilidad en la investigación naturalista, en: GIMENO, J. y PEREZ, A.: *La enseñanza: su teoría y su práctica*. Madrid. Akal Editor, 148-165.
- GUBA, E. G. y LINCOLN, Y. S. (1990): *Fourth Generation Evaluation*. London. Sage Publications.
- HUBERMAN, A. M. y MILES, M. B. (1991): *Analyse des données qualitatives. Recueil de nouvelles méthodes*. Bruxelles. ERPI/De Boeck Université.
- LOZANO, A. (1992): *El profesorado de Educación Compensatoria: Análisis de su discurso*. Madrid. UNED. Facultad de Ciencias de la Educación. (Tesis Doctoral inédita).
- MASTERMAN, M. (1975): La naturaleza del paradigma, en: LAKATOS, I. y MUSCRAVE, A.: *Crítica y conocimiento*. Barcelona. Grijalbo.
- MOLINA, S. (1992): La ideología pedagógica, en MOLINA, S. y GÓMEZ, A.: *Mitos e ideologías en la escolarización del niño deficiente mental*. Zaragoza. Mira Editores.
- MORGAN, D. L. (1988): *Focus Group as Qualitative Research*. London. Sage.
- MORGAN, G. (1990): *Images de l'Organisation*. Québec. Eska.
- POPKEWITZ, T. S. (1988): *Paradigma e ideología en investigación educativa*. Madrid. Mondador.
- SABIRON, F. (1992a): Referentes epistemológicos en evaluación etnográfica de centros docentes, en: III Colloque National «*Evaluation en Education*». Lisboa, 20-21 nov., Faculdade de Psicologia e de Ciências da Educação.
- SABIRON, F. (1992b): La culture de l'établissement scolaire. Vers une approche de l'analyse des Institutions Educatives, en XIV^o Congrès de l'AISLF, *Les Nouveaux Mondes et l'Europe*. Lyon. Université Lumière.
- SABIRON, F. (1993): Le processus d'amélioration sociale des établissements scolaires. Vers une démarche évaluative et ethnographique, en: *Por un nouveau bilan de la sociologie de l'éducation*. Paris. Institut National de Recherche Pédagogique.
- SABIRON, F. y LOZANO, A. (1990): El estudio de caso en la evaluación de centros escolares: Presentación y revisión de algunas dificultades, en: *Jornadas de Estudio sobre el Centro Educativo*. La Rábida. GID.
- SEGUIN, F. y CHANLAT, J. F. (1987): *L'analyse des organisations. Une anthologie sociologique*. Quebec. Gaëtan Morin éditeur.
- WEICK, K. (1983): «Educational Organizations as Loosely Coupled Systems». En: BALDRIDGE, J. V. y DEAL, T.: *The Dynamics of Organizational Change in Education*. Berkeley. Mc Cutchan Publishing Co, 15-37.

LA CONSTRUCCIÓN DEL CONOCIMIENTO EN LOS ESTUDIANTES UNIVERSITARIOS

Concepción Medrano Samaniego
Universidad del País Vasco

LA METACOGNICIÓN Y EL APRENDIZAJE SIGNIFICATIVO

Las personas que trabajamos como docentes en el ámbito universitario, sabemos que enseñar no es sólo informar acerca de principios o teorías que otros han elaborado, y también que contruir el conocimiento o aprender no es únicamente repetir lo que los profesores o textos nos muestran.

Las investigaciones llevadas a cabo en la última década en torno a la metacognición, nos hacen reflexionar sobre la necesidad de introducir cambios tanto en la forma de enseñar como en el modo de aprender. Sin embargo, ningún profesional de la enseñanza suele dudar o cambiar sus métodos para transmitir los conocimientos estructurados y sistematizados de la disciplina que imparte, mientras crea que son los mas adecuados o idóneos en sus actuales circunstancias.

Burón (1993) ha definido el objeto de estudio de la metacognición en los siguientes términos: es el conocimiento de las distintas operaciones mentales y saber cómo, cuando, y para qué debemos usarlas. Si desarrollamos la capacidad para reflexionar sobre lo que hacemos, estudiamos, practicamos etc.. en los alumnos universitarios, podemos establecer un estilo de enseñanza que favorezca el aprendizaje significativo y la toma de conciencia como estrategias para la construcción del conocimiento.

En el transcurso de mi experiencia como docente, siempre he sentido un conflicto entre transmitir conocimientos y dedicar «espacios» para la discusión, reflexión, cambio de esquemas conceptuales, transferencia de los aprendido a situaciones reales etc... en definitiva a cómo conjugar el viejo tópico de aprender a aprender y la trasmisión del conocimiento elaborado y estructurado.

En la experiencia que se presenta se ha pretendido:

— Que los estudiantes posean una idea clara del significado y alcance de la asignatura (psicología de la educación).

— Que conozcan las dificultades de una definición epistemológica y conceptual de dicha disciplina.

— Que ejerciten la reflexión acerca de los conocimientos elaborados y estructurados que se imparten en la enseñanza universitaria.

— Que utilicen las estrategias de autoregulación del pensamiento como favorecedoras de la comprensión frente a la mecanización.

LOS CONOCIMIENTOS PREVIOS Y LA INSTRUCCIÓN FORMAL

Por otro lado los estudiantes poseen conocimientos previos, es decir, construcciones personales que se han elaborado en el transcurso de su interacción con el medio ambiente, y que en muchas ocasiones prevalecen después de la instrucción formal.

Pienso que en muchas ocasiones a pesar de movernos en marcos teóricos constructivistas e intentar comunicar en nuestra docencia la importancia que tiene el que el propio sujeto sea el que construye sus aprendizajes, no hemos tenido demasiado en cuenta los «esquemas de conocimiento» que nuestros propios estudiantes universitarios manejan, por lo que no se ha considerado una de las condiciones para la consecución de aprendizajes significativos.

Diversos autores (Coll 1987; Driver, Guesne y Tiberghien 1989; Pozo y otros 1991), han señalado las características de los conocimientos previos y de los «esquemas» que los estudiantes utilizan para comprender el mundo que les rodea. Estas investigaciones han destacado la necesidad de que los estudiantes tomen conciencia de sus propias ideas incluso que las hagan explícitas. Para poder modificar nuestros «esquemas de conocimiento», adaptándolos al conocimiento científico, tenemos que ser conscientes de nuestras percepciones erróneas.

Desde una perspectiva constructivista, el aprendizaje significativo implica un cambio de esquemas de conocimiento. Y esto también implica un cambio metodológico en nuestra manera de transmitir «los contenidos curriculares». Este cambio supone descubrir las representaciones espontáneas de los alumnos y alumnas antes de transmitir los conceptos elaborados.

Trabajar desde una perspectiva constructivista supone, entre otras muchos aspectos concebir las estructuras cognitivas de nuestros estudiantes como un conjunto de esquemas de conocimiento. Estos esquemas pueden acercarse «más o menos» a la interpretación de la realidad, pero es importante no ignorarlos si queremos producir conocimientos contruidos versus conocimientos mecanizados o repetitivos.

Teniendo en cuenta la importancia de conocer los conocimientos previos y la reflexión acerca de aquello sobre lo que aprendemos, me interesé por indagar a través de un cuestionario semiestructurado (siguiendo la técnica de la entrevista clínica) las representaciones espontáneas que los estudiantes poseían en torno a la asignatura que imparto. En concreto, se trataba de que se argumentara en torno a la disciplina de «psicología de la educación» que en el Plan de estudios vigente se incardina en cuarto curso de la actual licenciatura de Pedagogía (Universidad del País Vasco).

De alguna manera las concepciones erróneas hay que entenderlas como el desequilibrio cognitivo necesario para la consecución de mayores cotas en la construcción del conocimiento.

PROCEDIMIENTO Y RESULTADOS

Este cuestionario escrito se ha pasado a 120 estudiantes y a través de las preguntas que se incluían en él, se pretendía acceder no sólo a la representación de qué es para ellos la «psicología de la educación», sino también las expectativas, intereses y alcance que para ellos suponía la propia asignatura.

Se aplicó durante la primera semana de clase con el fin de:

— Adecuar las ideas previas que poseen a lo que en realidad se va a trabajar en psicología de la educación.

— Reorganizar en lo posible los intereses que han manifestado y proporcionar información para la reflexión y discusión sobre el alcance y los límites de la asignatura de una manera realista.

Este cuestionario es de lápiz y papel y se realiza aproximadamente en 45 minutos. Los estudiantes que no pudieron contestar al cuestionario (14 en este estudio), lo hicieron oralmente en las horas de

tutoría. El hecho de aplicarlo oralmente permite profundizar más cualitativamente en algunos aspectos del cuestionario. Sin embargo, en cuanto al contenido no se obtiene una mayor información que en los cuestionarios aplicados por escrito. Las edades de los estudiantes van desde los 22 años del grupo diurno, hasta los 40 años del grupo nocturno.

Después de realizar el análisis de contenido de las respuestas obtenidas en torno a las ideas previas de la disciplina de «psicología de la educación» nos encontramos con los siguientes resultados:

De la muestra total, 62 estudiantes conciben psicología de la educación como una psicología evolutiva aplicada a la educación.

Aproximadamente la mitad de las alumnas y alumnos se representan la asignatura como una continuación de la psicología evolutiva pero adaptándola a las funciones que ellos consideran que debe desempeñar un profesional de la educación.

Casi una tercera parte de los estudiantes, 30 sujetos, han explicado sus ideas previas, en lo que se puede agrupar como una psicología del aprendizaje aplicada a la educación.

Otros 18 estudiantes, entienden la asignatura como una psicología de la instrucción, haciendo hincapié en la fundamentación teórica para el desarrollo del currículum.

Los 10 estudiantes restantes, presentan concepciones muy dispares que resultan difíciles de agrupar bajo un mismo concepto. Desde el que piensa que es el estudio de las variables psicológicas y/o personales hasta el que reduce la psicología al estudio de el psicoanálisis aplicado a la educación, o el que piensa que se trata de conocer las técnicas psicológicas para programar una educación de calidad.

Resumiendo, todas estas respuestas nos hacen reflexionar acerca de la dificultad de encontrar una concepción en psicología de la educación que unifique criterios y que otorgue a la propia disciplina una identidad propia.

REFLEXIONES FINALES

El hecho de poder reflexionar acerca de las concepciones de los estudiantes respecto a la asignatura, así como de las posibilidades y límites de la misma, ha sido en sí mismo un aspecto motivacional para desarrollar estrategias de enseñanza desde un modelo constructivo. Por ejemplo, es importante que los alumnos tomen conciencia desde el comienzo del curso que no existe un modelo global que explique el aprendizaje humano.

Al reflexionar acerca del significado de la propia asignatura, se logra no crear falsas expectativas, a la vez que se fomenta la necesidad de continuar estudiando e investigando.

Para favorecer la construcción del conocimiento es importante indagar en los conocimientos previos y elaborar conjuntamente con los estudiantes a través de la propia instrucción interpretaciones más adecuadas de la realidad.

Las estructuras de conocimiento pueden ser modificadas por la experiencia y la exposición o transmisión de la información, en la medida que se asimila la nueva información se modifican nuestras estructuras. En este sentido los esquemas son procesos activos que tienden a adaptarse a los datos que proceden del medio ambiente. En cierta medida un aprendizaje eficaz dependerá de la activación y reestructuración de los esquemas existentes.

Es importante señalar la similitud que se puede observar en la evolución de las ideas previas de los estudiantes con la efectuada en otras épocas en las mismas áreas de conocimiento. Este hecho se está utilizando como punto de referencia para la organización y secuenciación de contenidos.

BIBLIOGRAFÍA

AUSUBEL, D. P. (1982): *Psicología educativa*. México, Trillas.

- BANDURA, A. (1982): *Teoría del aprendizaje social*. Espasa-Calpe, Madrid.
- BURÓN, J. (1993): *Enseñar a pensar*. Mensajero, Bilbao.
- CARRETERO, M., POZO, J. I., y ASENSIO, M. (1989): *La enseñanza de las Ciencias Sociales*. Visor, Madrid.
- COLL, C. (1987): *Psicología y currículum*. Laia, Barcelona.
- COLL, C. (1988): *Conocimiento psicológico y práctica educativa*. Barcanova, Barcelona.
- DRIVER, R., GUESNE, E. y TIBERGHIE, A. (1989): *Ideas científicas en la infancia y en la adolescencia*. Morata, Madrid.
- GIORDAN, A. (1987): *Les origines du Savoir*. Delachaux et Niestlé, Neuchatel.
- MEDRANO, C., MUNARRIZ, B., MARTÍNEZ, B. y BASURKO, F. (1990): «Un modelo de intervención basado en la génesis de los conceptos y en la interacción social en las enseñanzas medias». *Revista de Investigación Educativa*, 8,16, 351-355.
- PORLÁN, R. y MARTÍN, J. (1991): *El diario del profesor*. Diada Editoras, Sevilla.
- POZO, J. I., SANZ, A., GÓMEZ CRESPO, M. A. y LIMÓN, M. (1991): «Las ideas de los alumnos sobre la ciencia: una interpretación desde la psicología cognitiva». *Enseñanza de las Ciencias*, 9, 1.
- SASTRE, G. y MORENO, M. (1980): *Descubrimiento y construcción de conocimientos: Una experiencia de pedagogía operatoria*. Gedisa, Barcelona.

EL ESTUDIO DEL AUTOCONCEPTO EN CHICOS Y CHICAS CON TÉCNICAS CUALITATIVAS

M^a Ángeles de la Caba Collado
Universidad del País Vasco

El **objetivo** es analizar el autoconcepto con el propósito de establecer, en la medida de lo posible, diferencias evolutivas y de género, en cuanto a sus aspectos centrales.

Más concretamente se pretende indagar sobre las siguientes cuestiones:

- 1) — Cuáles son los datos físicos en que me más se fijan para autodescribirse.
- 2) — Qué actividades, gustos, intereses y preferencias se manifiestan.
- 3) — Cuáles son las personas significativas más importantes a los que se hace referencia y forma de situarse.
- 4) — Qué datos psicológicos (carácter, forma de ser) se subrayan.
- 5) — Qué valoraciones sobre las propias habilidades se recogen tanto desde uno/a mismo/a como desde las personas significativas.

En cuanto a la muestra se han tomado un total de 160 redacciones, desde primero de EGB, teniendo en cuenta que el reparto de chicos y chicas fuera similar.

- 24 de Ciclo Inicial
- 68 de Ciclo Medio
- 68 de Ciclo Superior

COMO metodología se ha empleado la técnica de análisis de contenido, adecuada para analizar y describir, detalladamente, los datos que se aportan en un texto escrito en forma de redacción libre (Trend 1986 IEPS 1979).

Las **unidades de análisis** se han confeccionado haciendo un largo listado de todos los elementos que aparecen, ordenados según temas. Para ello se han tenido en cuenta los hallazgos de la Psicología Evolutiva. Básicamente se han considerado cuatro grandes temáticas propuestas por Damon (1979): autoconcepto físico, activo, social y psicológico. Además se han considerado otras áreas temáticas, como el «yo ideal» y la «autovaloraciones» (Hidalgo y Palacios 1992).

La **cuantificación**, como es habitual en esta técnica, se basa en el recuento de frecuencias.

DISCUSIÓN DE LOS DATOS

Ciclo inicial

Hay un claro predominio, en la línea propuesta por Damon, del autoconcepto físico y activo. Tanto niños como niñas tienen a describirse en función de sus rasgos corporales (pelo, ojos, altura, gordura),

sus ropas, los años que tienen, el lugar donde viven y en función de las actividades que realizan. Aparece, frecuentemente, la asistencia a la escuela, el juego y las salidas con los padres o viajes. Asimismo, los gustos y preferencias se relacionan con estas actividades. Destaca, especialmente, el gusto por ir a la escuela y lo relacionado con ella (pintar, escribir, leer). Se hace alusión, también, al gusto por jugar, por viajar y se mencionan platos preferidos. En el caso de los chicos aparece el fútbol, como hobby.

En cuanto a la importancia de los otros significativos, aparecen los padres y hermanos, a los que también se describe en términos físicos, destacando, como en su propio caso, el nombre o cosas que hacen juntos. «mi mamá se llama»... Si se menciona a los compañeros es para decir que se juega con ellos. Las menciones de la amistad son muy generales y se limitan a afirmaciones del tipo «tengo muchos amigos» o «mi mejor amigo-a es.»

Se trata, en general, de un período de satisfacción y bienestar, tanto con uno mismo como hacia los demás y el entorno. En el primer sentido, destacar frases del tipo «soy feliz», «vivo feliz». En cuanto al entorno: «este mundo es genial», «mi familia es muy buena».

Las autovaloraciones psicológicas, del carácter, son muy generales y o bien tienen que ver con las habilidades del juego— «se saltar muy bien' «juego muy bien al fútbol»!, o bien a las habilidades escolares «dibujo muy bien», «escribo muy bien», «leo muy bien», o bien a las relaciones inmediatas «soy obediente», soy travieso»...La única autoreferencia personal que no va en esta línea pone, una vez más, de manifiesto el sentido positivo de esta etapa: «Soy alegre, «soy divertida».

En este ciclo no hemos encontrado ninguna diferencia significativa en función del sexo.

Ciclo medio

En líneas generales, no hay grandes diferencias con respecto al período anterior y sigue predominando la autodefinición en términos físicos y activos. Sin embargo, hay una mayor diferenciación en todos los terrenos. En cuanto a las actividades, la escuela sigue siendo, un elemento de referencia para expresar lo que a cada cual le gusta. Por otra parte, destacar la televisión como elemento en torno al cual se articulan gustos y preferencias «mi programa favorito»...

Se esboza, clara y progresivamente, un «Yo ideal», tanto en chicos como chicas, en el cual destaca lo que gustaría hacer. Frente a la exaltación del hogar y el entorno, que había en el período anterior: «ir a Disneylandia», «vivir en Hawai o en una isla», «ser rica y tener un chalet con piscina», «ser famoso» aparecen con cierta frecuencia.

La autoestima, positiva o negativa, va cobrando gran fuerza y está muy relacionada con la escuela «voy bien-mal en el cole», «soy bueno en» y sospechamos que con ciertos hábitos de higiene. Aparecen junto a consideraciones de soy «sucio», «mentiroso», otras como «soy malo». Se ve, claramente, el peso que tiene lo que se recibe de los otros en la formación de esta autoestima. «dicen que soy majo», «dicen que soy guapa» (28%).

En cuanto al autoconcepo social la familia sigue ocupando un papel privilegiado, sobre todo los hermanos, con los que se puede jugar y los compañeros o amigos van cobrando fuerza «paso el tiempo con» «me lo paso bien con», «mi mejor amigo es» «tengo muchos amigos» y aparecen las consideraciones de los niños sobre cómo deben de ser las relaciones: «juego con todos», «dejo las cosas a los demás».

En este ciclo se empiezan a constatar importantes diferencias en función del sexo.

Aparecen diferencias claramente constatables según el sexo en cuanto a los gustos y disgustos. Se constata, en este sentido, una mayor preferencia de los chicos hacia las Matemáticas y de las chicas hacia la Naturaleza o las Ciencias sociales.

Cabe destacar, en el caso de las chicas, una mayor aparición de rasgos físicos relacionados con las manos, los dedos y las uñas.

Asimismo llama la atención, la importancia que va cobrando, la diferenciación de actividades. En el deporte, que se cita junto con la música como actividades favoritas, el fútbol, es en los chicos, la estrella (90%) frente a otros deportes como el baloncesto, la gimnasia y la natación en las chicas. Por otra parte, el uso del ordenador, el gusto por «ver consolas» parece, también, muy habitual en chicos (25%) pero no aparece en el caso de las chicas. Especialmente destaca, en el caso de éstas el gusto por estudiar (40%) mientras que en los chicos se cita mucho menos y en muchas ocasiones como la actividad menos preferida (35%). «lo que menos me gusta es ir al cole» «no me gusta trabajar», «lo que no me gusta es hacer los deberes» Entre las cosas que las chicas citan como de poco agrado están las siguientes «no me gusta ver a la gente triste», «no me gusta dejar sola a la gente porque me da pena», «me gustaría que no hubiera racismo».

También de forma diferenciada, destacar la alusión a la belleza en el caso, sobre todo de las chicas «soy guapa» o «tengo un pelo muy bonito» mientras que, curiosamente, la fuerza aparece como motivo de orgullo en los chicos. En este sentido, destacar que en los dos o tres casos de chicos que hacen alusión a la belleza la califican como «estoy macizo» y va unido a «soy un ligón» y a una clara referencia a la clase de chica que les gusta.

En los rasgos del carácter que se atribuyen también se encuentran diferencias. De hecho, nos ha llamado mucho la atención el altísimo porcentaje de chicas que se autodefinen como «trabajadora» o «estudiosa», frente al considerable porcentaje de valoraciones de los chicos del tipo «trasto», «peleón» «trasto» «gamberro» «chapucerete». Otros atributos como «simpático-a», «alegre» tímido-a» «listo-a» se distribuyen de forma similar.

Ciclo superior

Hay una clara diferencia con respecto a los dos ciclos anteriores y radica en el peso clave que juega el autoconcepto social y las claves autoreflexivas sobre uno mismo.

La referencia a los aspectos físicos disminuyen en los chicos mientras que parecen seguir siendo importantes en el caso de las chicas (20% frente a 80%).

En cuanto a las actividades, gustos e intereses la escuela pierde su anterior peso para dejar paso a otros aspectos. Cuando se cita a la escuela es, prácticamente siempre, para hablar del «deber» y la «necesidad» de estudiar para obtener trabajo y posición de cara al futuro. Por otra parte, la tendencia a favor de las chicas en cuanto su definición e identificación con los trabajos de la escuela va desapareciendo «soy una estudiante ni buena ni mala, regular y un poco más» «soy buena estudiante aunque es raro en las chicas de mi edad». «odio estudiar aunque algunas me consideren «empollona». Si estudio es porque sé que es mi deber».

Entre los intereses de este período hay que destacar dos grandes bloques.

Por una parte, están los hobbies entre los cuales el deporte y la música continúan teniendo un papel clave, aunque se nota la presencia de nuevos gustos y aficciones «me gusta ir a la discoteca» «me gusta fumar y beber».

Por otra parte, están los fortísimos intereses sociales: «me gusta conocer gente» «me gusta dar buena impresión». «no me gusta ser egoísta». En este sentido, cabe destacar que esta tendencia, si bien es general, está mucho más acentuada en las chicas «soy una chica que si puede te ayuda en todo», «soy una chica a la que no me gusta reír de nadie y ayudar» (60%).

La importancia de los compañeros y amigos como otros significativos es evidente «tengo muchos amigos y éso será por algo» «mis amigos dicen que les caigo bien, eso se deberá a que soy simpático, agradable y cariñoso» «dicen que soy muy majo. «me gusta que me aprecien»

La creciente capacidad lógica es manifiesta y evidente en la forma de autoanalizarse.

Así por ejemplo, se separa, claramente, lo bueno y lo malo que uno tiene en la autodefinición «soy

listo pero me considero feo», «soy majo pero tengo cosas mala. Soy despistado y no hay día que no me olvide algo» Variabilidad de ánimo «soy alegre pero a veces estoy triste». «a veces tengo mal humor pero se me pasa en seguida».

Aparece la referencia a los sentimientos y la comunicación (mucho más en las chicas) la preocupación a lo que piensan otros «me preocupa un poco ese ¿qué dirá la gente de mí? pero me preocupa más saber ser yo misma respecto a los demás.

En la relación con los otros hay un aspecto de carácter y personalidad que parece claramente diferenciador, mientras que otros no lo son, entre chicos y chicas. Ser un tío «legal» para los chicos y ser «sensible» para las chicas.

Cabe destacar, asimismo, la aparición de un **mundo interior** al que se alude a menudo «por dentro estoy sola».

Finalmente, hacer mención de los ideales, entre los cuales destacan por orden de frecuencia los siguientes: aspiraciones de riqueza y bienestar, la naturaleza, un mundo de igualdad y sin racismo.

REFERENCIAS BIBLIOGRÁFICAS

DAMON, W. (1979): *The Social World of the Child*. San Francisco: Jossey.

IEPS (1988): *Técnicas para analizar la solidaridad en la escuela*. Madrid: Narcea.

PALACIOS, J., MARCHESI, A., COLL, C. (1992): *Desarrollo psicológico y Educación*. Madrid: Alianza Psicología. (1990).

TREND, M. G. (1986): «Sobre la reconciliación de los análisis cualitativos y cuantitativos» en Cook, T. D. (eds.): *Métodos cualitativos y cuantitativos en investigación educativa*. Madrid: Morata.

ESTUDIO DE NECESIDADES PARA UNA ADAPTACIÓN CURRICULAR

M^a Celia Isorna Sánchez, Rosa Morillo Peña, M^a José Vecino Jiménez
Universidad de Sevilla

DESCRIPCIÓN DEL PROBLEMA

Nos planteamos el conseguir un programa adaptado a las Necesidades de un niño de ocho años, diagnosticado con disfunción cerebral mínima, en su variedad de síndrome hiperquinético.

Hemos considerado oportuno llevar a cabo un «estudio de caso» cuya definición recogemos de Yin (1984) definido como «el **análisis detallado de una situación, acontecimiento, sujeto o conjunto de documentos**, para lo cual se maneja múltiples fuentes de evidencia. Como método de investigación se utiliza para **profundizar en el conocimiento de los fenómenos sociales complejos** (descripción, análisis e interpretación), sean estos individuales, organizativos, sociales o políticos» (Álvarez, 1993).

Dentro del estudio de casos se ha optado por el tipo «estudio de caso único» que se centra en algo muy específico como puede ser un síndrome o trastorno.

METODOLOGÍA

«La investigación de tipo cualitativo, y en particular el estudio de caso único, trata de ser lo más riguroso y exhaustivo posible» (Oliver, 1993) estas son palabras que nos sirven de base y justifican nuestro posicionamiento.

Basándonos en Ibáñez (1990) se puede declarar que la investigación, desde una perspectiva cualitativa, se justifica a tres niveles:

- 1.— Nivel epistemológico, es decir, para qué o para quién se hace la investigación.
- 2.— Nivel metodológico o, lo que es lo mismo, por qué se hace así
- 3.— Nivel tecnológico, es decir, cómo se hace.

TÉCNICAS DE RECOGIDA Y ANÁLISIS DE DATOS

El Plan de Análisis elegido por el equipo de nuestra investigación está basado en autores como García Jiménez (1992), Bardín (1986) y Taylor y Bogdan (1986). Revisando la información que nos han aportado estos autores se ha creado el esquema que sigue:

- 1.— Lectura individual y grupal de los datos seleccionados

2.— Selección de técnicas de análisis de datos:

- La observación de conductas
- Entrevistas a padres y profesores
- Análisis de contenido de documentos (informes de estudios anteriormente realizados por médicos y psicólogos).

3.— Triangulación: confrontar los datos recogidos por los distintos miembros del grupo.

Nos han parecido más idóneas estas técnicas para lo que se pretende hacer, que es deducir las necesidades educativas y conductuales particulares del niño y formular un programa para satisfacerlas, basándonos en palabras de Hammill (1982).

También se ha utilizado una prueba estandarizada (Bücher) para hacer un primer acercamiento al problema.

El método de registro ha sido la toma de notas de campo (confección de un diario de campo) y la grabación mediante sistemas de audio.

El **análisis** de los datos recogidos se ha basado en el *análisis de contenido* de cada una de las entrevistas, a partir de las cuales se han extraído los siguientes puntos claves:

- * Ambiente familiar y escolar
- * Comportamiento del niño en los distintos contextos
- * Organización del centro
- * Disponibilidad de recursos...etc

El análisis del diario de campo que ha sido utilizado para contrastar y complementar toda la información adquirida.

El análisis del test ha sido realizado mediante una interpretación del dibujo hecho por el niño.

CONCLUSIONES

A partir de los datos arrojados, llegamos a la conclusión de que la adaptación curricular más idónea para este caso es una adaptación curricular significativa, entendiéndose por adaptaciones curriculares significativas: «aquellas que consisten principalmente en la eliminación de contenidos esenciales o nucleares y/u objetivos generales que se consideran básicos en las diferentes áreas curriculares y la consiguiente modificación de los respectivos criterios de evaluación» (Guía de adaptaciones curriculares, 1992).

Para la adaptación curricular son necesarios una serie de modificaciones en muchos sentidos:

- A.— a nivel metodológico
- B.— a nivel de objetivos
- C.— a nivel de contenidos
- D.— a nivel organizativo
- E.— a nivel de evaluación

A. Nivel Metodológico

Dentro de éste nivel hay que tener en cuenta:

- * Participación activa del alumno (actividades grupales...)
- * Motivación extrínseca e intrínseca en el proceso enseñanza-aprendizaje...etc.

B. Nivel de Objetivos

En ésta adaptación curricular ha sido necesaria la eliminación de objetivos a conseguir por el

alumno. De la relación de objetivos que el profesor marca para los alumnos se han eliminado una gran mayoría y otros han sido sustituidos por otro tipo de objetivos que el niño pueda alcanzar.

Entre los objetivos a conseguir se encuentran los siguientes:

- * Autonomía personal y social.
- * Aceptación y respeto hacia las diferencias individuales.
- * Participación en actividades grupales respetando las normas y reglas de convivencia
- * Observar y diferenciar características visuales y de relación espacial significativas en la percepción de la realidad cotidiana

C. Nivel de Contenidos

Los contenidos han de estar organizados en temas vivenciales, en centros de interés que tengan significación para el niño. Este aspecto es requisito imprescindible para el interés del niño, pero no sólo para éste niño sino para todos sus compañeros. Por eso, los contenidos no tienen por qué ser sustituidos sino adaptados.

D. Nivel Organizativo

Dentro de este nivel son necesarios los siguientes cambios para llegar a una Socialización:

- * Coordinación del grupo (colaboración entre el profesor de apoyo el tutor, el director y especialistas del centro).
- * Estructuración de tiempos.
- * Estructuración del aula, entre otras.

E. Nivel Evaluativo

«La Evaluación se plantea no como un juicio o valoración simplificadora del tipo «aprobado-suspenso», sino como elemento orientador; como una concepción de la evaluación como proceso continuo, simultáneo al proceso de enseñanza-aprendizaje» (Hervás, 1993).

Se puede decir que la evaluación está en función de las habilidades y destrezas que el niño desarrolle para poder conseguir muchos de los objetivos que éste puede alcanzar.

Además, la evaluación ha de ser formativa, continua y global, lo que supone la recogida de datos, con el fin de obtener la información necesaria sobre los procesos de enseñanza-aprendizaje que tienen lugar en el aula.

BIBLIOGRAFÍA

- ÁLVAREZ, V. B. (1993): «Los Procesos de intervención orientadora en los equipos de promoción y orientación educativa (EPOEs). Estudio etnográfico de un caso». *Proyecto docente* (Inédita).
- BARDIN, L. (1977): *Análisis de contenido* (Madrid, Akal).
- CONSERJERÍA DE EDUCACIÓN Y CIENCIA (1992): *Guía de adaptaciones curriculares*. Conserjería de Educación y Ciencias.
- GARCÍA JIMÉNEZ, E. (1991): *Una teoría práctica sobre la evaluación. Estudio etnográfico* (Sevilla, MIDO).
- HERVÁS, M. J. (1993): «Castilla y León. Una alternativa para la escuela en el medio rural». *Cuadernos de Pedagogía*, (211, 44-49).

- IBÁÑEZ, J. (1990): *Perspectiva de la investigación social: en el diseño en las tres perspectivas*. En García Ferrando, M. y otros (cop.): *Análisis de la realidad social. Métodos y técnicas de investigación* (Madrid, Alianza).
- OLIVER VERA, C. (1993): «El agrupamiento flexible». *Cuadernos de Pedagogía*, (212, 19-21).
- TAYLOR y BOGDAN, (1986): *Introducción a los métodos cualitativos de investigación* (Barcelona, Paidós).

SIMPOSIUMS

SIMPOSIUM I

¿ORIENTACIÓN POR PROGRAMAS O POR SERVICIOS?

M^a Teresa Díaz Allué,
Coordinadora

*Víctor Álvarez Rojo, Benito Echeverría Samanes,
Josefina Hernández Fernández, Francisca Arbe*
Participantes

LA INTERVENCIÓN POR PROGRAMAS: ¿UNA SIMPLE INNOVACIÓN TECNOLÓGICA?

Víctor Álvarez Rojo

CAMBIOS EN LA EDUCACIÓN E INNOVACIONES METODOLÓGICAS

La evolución de los sistemas educativos en los países de la OCDE, a tenor de las condiciones sociolaborales cambiantes, ha dado lugar a diferentes ofertas de asesoramiento a los centros educativos; éstas se han configurado como alternativas políticas a situaciones de crisis o frente a necesidades de innovación de la educación o simplemente de algunos niveles de enseñanza. Las dos modalidades organizativas básicas que se han utilizado para su configuración, y que han originado un debate dialéctico persistente, han sido las de *servicios de apoyo externo* al centro educativo vs. *intervenciones integradas* en el mismo. Probablemente, como ha ocurrido con otros enfrentamientos teóricos-aplicativos de sobra conocidos, las ventajas-inconvenientes de cada una de estas propuestas se deban a factores histórico-contextuales más que a la incompatibilidad intrínseca de ambas propuestas organizativas (Escudero, 1992)¹.

Refiriéndonos a nuestro país el modelo adoptado de forma generalizada por las diferentes comunidades autónomas ha sido el de servicios de apoyo externo, entre los cuales se encuentran los servicios de orientación. No obstante, esta opción inicial está sufriendo, en el ámbito de la orientación, una evolución perceptible hacia nuevas formas organizativas, imposible de calificar como positiva o negativa en cuanto a su incidencia real en el sistema educativo, debido a la falta de una suficiente perspectiva histórica. Por una parte, algunas comunidades autónomas (vg. Canarias. C.Valenciana, Navarra) han configurado los servicios de orientación con una voluntad más o menos explícita de

integración de los mismos en los centros educativos; esta tendencia ha recibido el respaldo de las propuestas de la LOGSE con el diseño de los dos primeros niveles de intervención orientadora (tutoría y departamento de orientación) integrados en los centros.

Pero además, desde hace aproximadamente una década se viene produciendo en España una apuesta, todavía casi exclusivamente a nivel teórico, por una alternativa de acción orientadora constituida por la *intervención por programas*. Las razones de esta innovación incipiente han sido, desde mi punto de vista, las siguientes.

a) El *fracaso* constatado, aun en el corto espacio de tiempo de existencia de los servicios de orientación, de las *intervenciones diagnóstico-remediales-externas*. Constatado por los propios orientadores cuyo rechazo a la función diagnóstica aparece en primer plano en las pocas pero significativas investigaciones realizadas sobre estos servicios en nuestro contexto (Álvarez, 1992; Escudero, 1992). Manejar la variable «número de diagnósticos realizados por curso académico» como un indicador de la calidad, eficacia o extensión de los servicios de orientación ya no se les ocurre hoy en día más que a los diseñadores de propaganda electoral.

Parecer ser una creencia compartida entre los profesionales de la orientación la falta de justificación de unos servicios cuyas intervenciones fuertemente polarizadas en el diagnóstico externo no tienen las mínimas garantías de continuidad ni en el aula ni en el centro y sirven con harta frecuencia para justificar a posteriori determinadas actuaciones profesionales, organizativas o burocráticas que consolidan, en vez de innovar, las formas de hacer de un determinado contexto educacional.

b) La *masificación* de los centros educativos frente a la *escasez de recursos materiales y humanos* de los servicios de orientación. Ha sido ésta una opción política cuyas consecuencias en la práctica profesional quizá se hayan minusvalorado. Podríamos decir, sin correr demasiado riesgo de equivocarnos, que ha sido la demanda social la que ha abonado el terreno para las innovaciones en la intervención, más que la «cruzada» de los teóricos de la orientación en pro de la intervención por programas. Lo que ha ocurrido es que, casi desde los comienzos de la actuación de los servicios institucionales de orientación, los orientadores se han visto presionados sistemáticamente para intervenir en «todas» las situaciones problemáticas y en «todos» los contextos (centros) educativos de sus zonas, tanto por parte de las administraciones educativas como por los propios agentes de la educación (profesores, primordialmente, y padres). Al ser ésta una demanda sencillamente imposible de satisfacer, los orientadores han cosechado injustamente el rechazo o las críticas de ineficacia que a otros correspondían en gran medida.

Este estado de cosas ha dado como consecuencia una búsqueda asistemática y bastante generalizada de mecanismos de supervivencia profesional: cómo intervenir de forma que un número significativo de los destinatarios potenciales de la orientación reciban algo cuando el número de orientadores no varía significativamente respecto a unas ratios de población y de problemas escolares a todas luces irracionales. Los programas comienzan a verse ahora como una posible alternativa.

c) La intervención del orientador como actividad *rutinaria, parcial escasamente interesante e inevaluable*. «¿Para qué sirve un orientador?», ha sido una pregunta que con frecuencia se han formulado muchos profesionales después de un cierto tiempo realizando diagnósticos cuyos resultados no parecen servir para gran cosa o son demandados por motivaciones latentes. Orientar ha significado durante bastante tiempo «diagnosticar alumnos de educación especial». Y la contemplación de los resultados de una batería de pruebas diagnósticas, repetida ad infinitum, no parece ser sustancialmente más motivante que apretar el mismo tornillo durante horas y días en una cadena de montaje de triste recuerdo en la organización industrial. Además, ¿qué claves de retroalimentación profesional proporcionan uno o cien diagnósticos? No obstante, en la universidad, en las obras de consulta, en los congresos... se insiste en que orientar es algo más.

Es necesario, pues, catalogar el aburrimiento de los propios profesionales de la orientación ante determinadas funciones que les han sido asignadas (y, sin duda ninguna, ante el diagnóstico) como un factor potencialmente innovador de considerable magnitud. Los programas se presentan como una forma de intervenir susceptible de vehicular los más variados contenidos educativos; contenidos (educación para la paz, prevención de drogodependencias, educación ambiental, educación para la carrera...) que, por otra parte, la sociedad progresivamente demanda en estos momentos a los orientadores y no a otros profesionales.

d) Una evaluación progresiva, efectuada por los propios orientadores al irse consolidando lo equipos, de los modelos de intervención utilizados hasta el momento. Dicotomizando éstos en modelos de intervención directa e indirecta, la estrategia de intervención deseable parecer ser, a tenor de las investigaciones ya citadas y de alguna otra (Álvarez, 1993)³, la contemplada por el *modelo de intervención indirecta* cuyos principios programáticos pueden sintetizarse así:

- La intervención en los centros exige un determinado grado de **colaboración / implicación del profesorado** para que tenga sentido y consecuencias en la realidad educativa del centro
- La actuación debe estar **integrada en la dinámica del centro**, no puede ser externa a la misma
- Cualquier intervención tiene que ser el producto de un **acuerdo de actuación orientadores-centro** sobre sectores, temas, metodologías..., etc.
- La intervención de los orientadores respecto a los destinatarios últimos de la intervención, los alumnos, ha de ser **preferentemente una actuación indirecta** (o, en su caso, mixta) a través de los profesores tutores.

HACIA UNA CORRECTA INTERPRETACIÓN DE LA INTERVENCIÓN POR PROGRAMAS

La intervención por programas se presenta en estos momentos como una tecnología apetecible y posibilista que, no obstante, no puede proponerse como la panacea de nada. Y requiere, además una correcta comprensión.

El modelo de intervención por programas surge, según ha señalado Gysbers (1981)⁴, en los primeros años de la década de los 70, aunque en la década anterior diversos movimientos (la orientación para la carrera, la orientación para el desarrollo, el desarrollo de la carrera, el movimiento en pro de la «accountability») ya habían planteado la necesidad de buscar alternativas a la orientación adaptativa, a las intervenciones terapéuticas y administrativas. Puede decirse que este modelo es una consecuencia necesaria de un movimiento general *en pro de la orientación para el desarrollo*. Este movimiento parece representar inicialmente una respuesta de supervivencia lanzada desde los ámbitos profesionales de la orientación, asfixiados, tanto en USA como en Europa, por la dinámica impuesta por el modelo de servicios (Cofsky, 1991; Landry, 1991)⁵.

Se presenta, por tanto, como un modelo de intervención mediante el cual el orientador puede ejercer funciones distintas a las del diagnóstico y la terapia y como una forma de intervención eficaz para hacer efectivos dos de los principios de la orientación educativa marginados en la práctica orientadora: el de prevención o proactividad y el de intervención social y educativa. La noción de programa lleva implícito un tipo de **intervención comprehensiva**, que atiende tanto a la prevención y al tratamiento remedial o terapéutico de situaciones problemáticas o crisis (cualesquiera que sean su naturaleza o contenido), como al desarrollo de situaciones educativas (Palmer, 1991)⁶.

Diferentes autores han llevado a cabo intentos de caracterización de los programas de orientación. En nuestro contexto cabe resaltar las aportaciones de Rodríguez Espinar (1986, 1992 y 1993)⁷ y de Sanz Oro (1990)⁸, entre otros, que aparecen en el Cuadro 1.

Las características de los programas han sido igualmente el objeto del trabajo realizado muy recientemente por Borders y Drury (1992)⁹. Se trata de un macro trabajo de revisión de las investiga-

CUADRO 1
 CARACTERÍSTICAS DE LA INTERVENCIÓN POR PROGRAMAS

RODRÍGUEZ ESPINAR	SANZ ORO
1) Encaminada a la consecución de objetivos globales de desarrollo de los alumnos 2) Integrando dicha intervención en las experiencias cotidianas de aprendizaje 3) En la que todos los agentes educativos participan, incluidos los orientadores, con funciones claramente delimitadas 4) Que exige la creación de un currículum propio de orientación (metas y actividades) 5) Para ser desarrollado a lo largo de toda la escolaridad 6) Evaluando los efectos/resultados de la intervención ejecutada 7) Y atendiendo las necesidades de formación para los propios ejecutores de la intervención	— La orientación se concibe como un programa educativo en el que están implicados no sólo los orientadores sino también los demás profesionales de la educación (profesores) y otros paraprofesionales (monitores..). — Persigue unos objetivos centrados en el desarrollo de competencias (y no en remediar déficits) mediante una intervención cuidadosamente planificada, ejecutada y evaluada. — El programa es una « propuesta de equipo »

ciones empíricas llevadas a cabo durante los últimos treinta años sobre programas de orientación, de las normas emanadas de las asociaciones profesionales de orientación en USA (AACD y ASCA) y de otras fuentes con objeto de obtener una caracterización de los programas. Las principales notas definitorias de los mismos aparecen a continuación.

— **Programas educativos independientes:** un programa ha de contar con un currículum propio, con metas y objetivos secuenciales y tareas / actividades de desarrollo; dirigido a los diferentes dominios del desarrollo del sujeto (madurez personal / emocional, escolar, para la carrera, madurez social) y para todos los niveles educativos; manejado por profesionales cualificados, con materiales didácticos propios y sometido a una evaluación sistemática.

— **Programas integrables,** es decir, ha de tratarse de una oferta de intervención que pueda funcionar, bien como componente independiente del programa educativo global de la institución o como parte integrada de dicho programa (= infusión curricular).

— **Programas para el desarrollo:** la intervención debe estar basada en alguna de las teorías del desarrollo humano, ser proactiva y dirigida a los alumnos que se enfrentan con las tareas normales de desarrollo en cada estadio evolutivo.

— **Programa equitativos:** han de diseñarse para todos los alumnos y, consecuentemente, englobar todas la peculiaridades de la población destinataria: desigualdades sociales o culturales, grupos étnicos, discapacitados y handicaps ambientales (rural / urbano, zonas deprimidas).

Estos autores consideran además de enorme importancia dos aspectos de la intervención con esta tecnología:

— La **revisión /renovación del programa:** se impone como exigencia para hacer frente a la transformación de los contextos y a las nuevas aportaciones teóricas sobre intervención en diferentes ámbitos.

— El **clima del programa:** está constituido por el tipo de intervenciones que ofrece (entre iguales, entre profesores / orientadores y alumnos), por el apoyo de las autoridades educativas, por las relaciones que ofrece establecer con los padres y la comunidad y por los recursos de que dispone.

Desde mi punto de vista, un programa de intervención puede *definirse* como la acción colectiva del equipo de orientadores, junto con otros miembros de la institución, para el diseño, implementación y evaluación de un plan (programa) destinado a la consecución de unos objetivos concretos en un medio socioeducativo en el que previamente se han determinado y priorizado las necesidades de intervención.

Es decir, un programa consiste en una acción intencional planificada que excluye la improvisación. Y es aquí donde reside el punto débil de esta innovación tecnológica en la medida en que se está utilizando el término programa para cualquier tipo de intervención que se lleva a cabo. Es más, da la impresión de que, en la práctica, se entiende por programa la parcelación del conjunto de tareas (o de las funciones) que los orientadores llevan a cabo, con una mayor racionalización de las mismas y la elaboración de algunos materiales de apoyo. No, un programa no puede ser un conjunto de actividades heterogéneas seleccionadas por criterios intuitivos o de simple práctica profesional («me ha resultado bien en otras ocasiones») (cuyo valor criterial, por otra parte, no estamos negando en estos momentos), por la sencilla razón de que la emisión de juicios de valor mediante la evaluación se convierte entonces en una tarea ardua: a qué atribuiremos entonces los resultados (positivos o negativos) de ese «programa», a la intuición del orientador, a sus habilidades personales, al clima del centro...; por qué esas actividades en el programa y no otras; cómo se justifica la organización y secuenciación del programa...

Así pues, una de las tareas fundamentales a la que ha de hacer frente el orientador dentro de este modelo de intervención es *el diseño de programas*. Es una tarea sistemática y no exenta de complejidad, de ahí que yo la haya considerado como el «talón de Aquiles» de este modelo de intervención, puesto que muchos orientadores consideran que no disponen ni del tiempo ni de los recursos de formación necesarios para llevarla a cabo. No obstante este handicap es perfectamente subsanable... si existe la voluntad política de potenciar los servicios de orientación.

Como dinámica de trabajo para abordar esta tarea, Kettner (Kettner y otros, 1990)¹⁰ propone, en un contexto amplio de trabajo social, la elaboración de programas de intervención de acuerdo al siguiente proceso:

- 1— **Determinación de las necesidades y de los problemas** de los destinatarios
- 2— **Selección de la estrategia de intervención** mediante la elaboración de hipótesis y la formulación de objetivos
- 3— **Diseño del programa de intervención** atendiendo a los siguientes elementos: recursos materiales y humanos, oferta de programa y resultados esperados

4— **Construcción de un sistema de información** sobre la efectividad del programa. La recogida de datos deberá realizarse en dos vertientes:

— Sobre lo que ocurre cuando el programa se pone en marcha: información para la toma de decisiones durante el proceso de implementación

— Sobre las características «estáticas» de los destinatarios, que probablemente influyen en la eficacia del programa (características sociales, educativas y «problemáticas»)

5— **Cálculo de Costos** de la intervención

6— **Planificación, administración y control del gasto**

7— **Evaluación** del funcionamiento del programa según cuatro categorías generales:

a) Evaluación del **esfuerzo**: Cantidad de actividades que se han llevado a cabo; cuáles se han realizado y cuáles no

b) Evaluación de **resultados**: en qué medida se han conseguido los objetivos del programa

c) Evaluación del **impacto** del programa: hasta qué punto el programa ha incidido en las necesidades del grupo (o comunidad) de destinatarios

d) Evaluación **costo/beneficio y costo/eficacia**: **costo por unidad de intervención y costo** para el logro de los resultados

No obstante, una consideración sintética de las diferentes propuestas sobre elaboración y aplicación de programas nos lleva a considerar la implicación del orientador en *cuatro tareas generales*:

1) Diagnóstico / Evaluación de las necesidades y/o problemas a los que se dirige el programa

2) Planificación y diseño del programa

3) Implementación del mismo en los contextos naturales

4) Evaluación de los diferentes elementos implicados en el diseño y aplicación, y revisión, en su caso, del programa («program review») una vez implantado de forma estable en uno o varios contextos

De estas cuatro tareas son suficientemente conocidos los procedimientos referidos al diagnóstico, implementación y evaluación (Pérez Juste, 1989¹¹, Colás, 1993¹²). Sin embargo, los intentos de sistematización de una metodología precisa para la planificación y diseño de programas ha sido más esporádicos. Cabe destacar en nuestro contexto las propuestas de Rodríguez Moreno (1992)¹³ y de Rodríguez Espinar y otros (1993). Desde mi propia perspectiva, el diseño de programas debe, en síntesis, proceder según aparece en el Cuadro 2.

MÁS ALLÁ DE LA TECNOLOGÍA

La intervención por programas supone una innovación tecnológica respecto a la intervención más o menos clásica derivada del modelo de servicios. Pero, ¿podemos hablar de un simple cambio tecnológico; es decir, de la sustitución automática de unas herramientas por otras más útiles en un momento o contexto dado? Responder afirmativamente al interrogante formulado sería tanto como dar por válida la insostenible pretensión de aquellos científicos (generalmente al servicio de alguien y de algo), que sostienen que la ciencia y la tecnología son apolíticas y no contaminadas ideológicamente.

No, no se trata de una simple innovación tecnológica. La asunción del modelo de intervención por programas lleva un mensaje político-ideológico implícito, cuyos postulados más sobresalientes son los siguientes.

— *La intervención por programas* es un modelo que responde a una *concepción diferente de la orientación*, basada no en el tratamiento exclusivo de problemas o crisis, sean éstos personales o grupales, sino en una reconceptualización del status de la actividad orientadora en el contexto institucional y ecológico de la educación. Responde igualmente a reconsideración de las funciones del orientador y a una revisión comprensiva y en clave de desarrollo de sus intervenciones. En el nuevo

CUADRO 2
 PROCESO PARA EL DISEÑO DE PROGRAMAS

1) ELABORACIÓN / SELECCIÓN DEL ESQUEMA CONCEPTUAL DEL PROGRAMA

- Selección de **teorías**: del desarrollo, del aprendizaje, de las relaciones con el entorno, de la relación de ayuda...
- Selección de la **propuesta curricular**: teoría curricular de base, metas educativas, propuestas de organización, criterios de evaluación
- Selección de **tecnología**:
De diseño: herramientas para acotar y configurar los diferentes elementos del programa (objetivos, contenidos, actividades...)
Instrumental: técnicas de ayuda aplicables (o bases para su construcción), técnicas de información, técnicas de evaluación y toma de decisiones, soportes didácticos

2) DELIMITACIÓN DE LA OFERTA GLOBAL DEL PROGRAMA

- Elaboración de hipótesis
- Definición del servicio que se quiere prestar
- Determinación de las metas y objetivos
- Fijación de las actuaciones requeridas por el programa
- Especificación del método de intervención (directo o indirecto) y del tipo de programa

3) DISEÑO DE LA OFERTA CONCRETA DE PROGRAMA

- Especificación de las unidades de intervención (selección y secuenciación de contenidos; construcción de sesiones de trabajo)
- Determinación de resultados esperados
- Fijación de la participación requerida de los destinatarios
- Identificación de recursos materiales y humanos

modelo éstas se centran en las necesidades y metas del contexto educativo y las ofertas de intervención se diseñan de forma programática y congruente con dichas necesidades y metas.

— Este modelo de intervención conlleva, implícita y explícitamente *una meta: la reconceptualización del status de la orientación*, tanto a nivel institucional como disciplinar. En el nivel institucional la orientación *reivindica una posición de igualdad* en el contexto curricular: el (o los) programa(s) de orientación, centrado en el desarrollo de diferentes ámbitos individuales y grupales y en la consecución de metas educativas debe tener igual posición y consideración que el programa instructivo. Se reivindica, por tanto, su inserción curricular para superar su carácter de «servicio de apoyo», «tapa-huecos», sustitutorio de determinadas deficiencias institucionales.

— Las consecuencias de este nuevo enfoque a nivel disciplinar son importantes. Por una parte, la orientación tiene que *reconceptualizar los contenidos curriculares de su oferta* educativo / orientadora en los diferentes ámbitos de intervención: desarrollo intelectual, social, institucional / organizativo y en relación con el contexto ecológico en el que los sujetos de la educación deben interactuar. No son válidos ya, por insuficientes, para esta perspectiva los contenidos tradicionales (trastornos de aprendi-

zaje, TTI, etc.). Además, la orientación se ve obligada a *poner a punto una metodología de intervención diferente* y adecuada a las exigencias del programa, para su diseño, implementación y evaluación.

— La determinación de *las funciones* de la orientación en este modelo de intervención está sujeta primordialmente a las *exigencias de elaboración y manejo de los programas*, en la medida en que la nueva «herramienta» de intervención lleva implícitos unos requerimientos, no sólo conceptuales/formales (como ya hemos visto), sino también funcionales: el orientador no es le «factor» único de las intervenciones orientadoras; éstas requieren como condición previa procesos de discusión, negociación y consenso para posibilitar su asunción y aplicación por parte de los agentes primarios de la educación (padres y profesores). Las funciones del orientador se desplazan, pues, desde la elaboración-aplicación-evaluación de acciones orientadoras hacia la dinamización-negociación-asesoramiento-apoyo a los procesos pactados de intervención.

Como afirmé inicialmente, la intervención por programas no es la panacea ni lo ha sido en ningún país. No puede ser contemplada como una estrategia excluyente (los servicios seguirán existiendo con sus funciones y prioridades políticamente asignadas) y su implantación requerirá algo más que la sustitución de una tecnología parcialmente ineficaz (exige un cambio de mentalidad respecto a qué es intervenir en la escuela).

Recién estamos empezando. Ya veremos que ocurre.

REFERENCIAS BIBLIOGRÁFICAS

- 1 ESCUDERO MUÑOZ, J. M., MORENO OLMEDILLA, J. M. (1992): *El asesoramiento a centros educativos (Estudio evaluativo de los equipos psicopedagógicos de la Comunidad de Madrid)* (Madrid, Consejería de Educación y Cultura).
- 2 ÁLVAREZ, V. y otros (1992): *La orientación institucional en Andalucía. Aportaciones para su evaluación* (Universidad de Sevilla, Grupo de Investigación MIDO).
- 3 ÁLVAREZ, V. (1993): *Los procesos de intervención orientadora en los Equipos de Promoción y Orientación Educativa (EPOEs). Estudio etnográfico de un caso* (Sevilla, inédita).
- 4 GYSBERS, N. C., MOORE, E. J. (1981): *Improving guidance programs* (Englewood Cliffs, Prentice-Hall).
- 5 COFSKY, L. (1991): «Les pièges relationnels de la profession au secondaire». *L'Orientation*, 5, 2, 23.
LANDRY, L. (1991): «Une image publique a battre». *L'Orientation*, 5, 2, 3.
- 6 PALMER, J. H., PAISLEY, P. O. (1991): «Student assistance programs: a response to substance abuse». *The School Counselor*, 38, 4, 287-293.
- 7 RODRÍGUEZ ESPINAR, S. (1986): *Proyecto docente e investigador* (Barcelona, inédita).
— (1992): «Orientación y reforma: el reto de la intervención por programas». *Currículum*, 5, 27-47.
— (Coord.) (1992): *Teoría y práctica de la orientación educativa* (Barcelona, PPU).
- 8 SANZ ORO, R. (1990): *Evaluación de programas en orientación educativa* (Madrid, Pirámide).
- 9 BORDERS, L. D.; DRURY, S. M. (1992): «Comprehensive school counseling programs: a review for policymakers and practitioners». *Journal of Counseling and Development*, 70, 2, 487-498.
- 10 KETTNER, P. M. y otros (1990): *Designing and managing Programs. An effectiveness-Based approach* (Newbury Park, Sage).
- 11 PÉREZ JUSTE, R., GARCÍA RAMOS, J. M. (1989): *Diagnóstico*.
- 12 COLAS BRAVO, P., REBOLLO CATALÁN, M. A. (1993): *Evaluación de programas. Un guía práctica* (Sevilla, Kronos).
- 13 RODRÍGUEZ MORENO, M. L. (1992): *Enseñar y aprender a tomar decisiones vocacionales* (Madrid, MEC).

Y, ¿LOS SERVICIOS SIRVEN?

Dr. Benito Echeverría Samanes
Universidad de Barcelona

¿LEY DEL PÉNDULO?

El debate actual sobre la intervención orientadora por «programas vs servicios» nos evoca la tesis mantenida por uno de nuestros más maduros profesores, que solía afirmar con buena dosis de ironía: La única ley vigente en educación es la del péndulo.

Al parecer, este artilugio no sólo sirvió a Foucault para demostrar la rotación de la tierra, sino que también algunos teóricos de la orientación tienden a aprovecharse de sus «oscilaciones con isocrónica majestad», perfectamente descritas por Umberto Eco. La influencia de la gravedad del péndulo les lleva a aferrarse a una de las puntas del hilo, que sustenta el cuerpo suspendido (= desarrollo de la autodeterminación de las personas), mientras que denostan con afán la otra.

Por «razones del guión» se nos ha pedido situarnos en la punta del sedal defensora de los servicios y aceptamos gustosamente el papel, siempre que se asuma de entrada el hecho comprobado de que el período de un péndulo no depende de la masa que esté oscilando, ni de su forma, ni de otra propiedad más que de la longitud del hilo, medida desde el punto de suspensión hasta el centro de gravedad del cuerpo que oscila.

Esta realidad es perfectamente adaptable al problema que nos ocupa y es esencial tenerla en cuenta, para poderle encontrar soluciones con perspectiva de futuro.

Si en este debate logramos acortar la longitud del hilo, probablemente se desvele el sofisma de que la intervención a través de servicios es únicamente reactiva y terapéutica, mientras que la realizada por programas tiene la exclusiva de la prevención y el desarrollo.

Esta tesis difícilmente se puede mantener, cuando se analizan comparativamente los modelos de intervención desarrollados en los estados de nuestro entorno más próximo, como hemos tenido ocasión de comprobar en fechas recientes (F. Company y B. Echeverría 1992).

LA REALIDAD SE IMPONE

Un análisis detenido de éste o de los otros informes sobre la «Estructura de los empleos y cualificaciones en el ámbito de la orientación escolar y profesional», promovidos por el CEDEFOP en los doce estados de la Unión Europea, puede ser esclarecedor al respecto, al permitir detectar aspectos tales como:

- La acción directa o indirecta de la intervención orientadora llega a muchos más beneficiarios en aquellos estados con mayor tradición de servicios de orientación. Dentro de la Unión Europea, el eje norte/sur marca sintomáticas diferencias a favor de los ciudadanos de los estados nortefños, donde la intervención, a través de servicios, está más arraigada.
- La escasez de recursos humanos y materiales explica en buena parte que en los estados mediterráneos se haya comenzado en los últimos años a intervenir por programas, más bien esporádicos y de corta duración, para paliar grandes déficits en algunos núcleos de la población, generalmente en los más desfavorecidos.
- Las citadas circunstancias condicionan en la mayoría de los casos los programas de orientación desarrollados, que, hoy por hoy, son predominantemente de carácter reactivo y sólo unos pocos llegan a la intervención social, con el fin aprovechar oportunidades de momentos históricos, que permitan modificar aspectos concretos del marco académico y socio-laboral.
- La funciones de prevención y desarrollo tienden a desarrollarse más en los estados que cuentan con servicios de orientación más consolidados y cada vez menos son terapéuticas, si se exceptúa el típico caso de Bélgica, donde históricamente la vertiente psico-médica ha gozado de amplia raigambre.
- Los servicios más estables han comenzado a concretar en bastantes casos las funciones de prevención y desarrollo a través de objetivos en conexión con el resto de servicios de sus respectivas comunidades, con el fin de responder directamente a las necesidades detectadas, tras previos análisis.
- La evaluación y seguimiento de las acciones realizadas paulatinamente vienen siendo asumidas por los servicios, primero por razones económicas, también por la presión que ejerce en algunos casos la corriente actual de «accountability» y aunque, en menor grado, por intereses de desarrollo profesional y perfeccionamiento de la práctica orientadora.
- La optimización de los recursos disponibles, exigida en tiempos de crisis, viene obligando a los servicios a desprenderse de grandes infraestructuras estáticas y a acercarse cada vez más a los entornos donde se generan las necesidades de los grupos beneficiarios de la intervención orientadora.

De una forma u otra estos indicadores dejan patente que la intervención por servicios puede colaborar al desarrollo de la orientación, tanto como la ejercida por programas y que las teóricas limitaciones, asignadas por algunos tratados a la primera, «per se» no son mayores ni más profundas que las que acompañan a las actuaciones llevadas a término desde el enfoque de programas.

Convendría recordar al respecto la interesante constatación de Umberto Eco, al recrear la experiencia de Foucault: «El péndulo reduce su velocidad en un extremo del plano de oscilación, para dejarse caer indolente hacia el centro y hundir el oculto cuadrilátero de fuerzas que marcan su destino».

En orientación estas fuerzas no son otras que sus cuatro principales funciones —terapéutica, preventiva, de desarrollo y de intervención social— y es un craso error situar el péndulo en un extremo y aprovecharse de la aceleración coyuntural, generada por su peso. No ha de olvidarse, que al abandonar un péndulo en un extremo adquiere una aceleración debida a su peso que le hace caer y que su velocidad va aumentando hasta llegar a la posición de equilibrio, precisamente cuando el hilo está en vertical.

Hacia este punto punto de equilibrio parecen tender hoy los sistemas de orientación de nuestro entorno más próximo, acercándose los enfoques del eje norte-europeo a las ricas experiencias de la cuenca mediterránea. Es decir, el modelo de servicios actuando por programas comienza a dar los frutos más provechosos.

PROGRAMAS A TRAVÉS DE SERVICIOS

La explotación de esta corriente de actuación aún es más necesaria en un país como el nuestro, que tradicionalmente ha contado con pocos servicios de orientación, desvinculados de la vida comunitaria y educativa, con pluralidad de competencias y compartimentalizados, faltos de planificación y coordinación, con situaciones financieras marginales y basados más en la buena voluntad de las personas, que en la consideración técnica de los problemas.

Este es un hecho aceptado por la mayoría de sus analistas, pero no compartimos la idea de que «en perro flaco, todo son pulgas». Más bien somos partidarios de alimentarlo y proclives a valorar en sus justos términos la infraestructura, que pueden ofrecer los servicios, para poder llevar a cabo programas de intervención orientadora con perspectiva de futuro.

Esta es al menos la experiencia contrastada en los últimos años desde el Servicio de Orientación Universitaria (SOU) de la División de Ciencias de la Educación de la Universidad de Barcelona.

En la segunda mitad de la década pasada fueron varios los programas de orientación puestos en marcha dentro de este entorno, pero su continuidad estuvo siempre afectada por la financiación de los mismos y la disponibilidad temporal de sus responsables y dinamizadores. Prácticamente todos respondieron bien a las necesidades de los universitarios, fueron altamente valorados por ellos y por los responsables de la institución, pero..., en el fondo estaban considerados más como dones gratificables, que como derechos y deberes a cumplir.

Hubo de cambiar previamente esta apreciación, para que desde 1991 sea posible venir desarrollando similares programas a los anteriores, además de otros derivados de la detección de importantes necesidades actuales, relacionadas con la elección de itinerarios formativos, preparación para la inserción laboral, etc.

Hoy en día, pocas personas cuestionan la función asignada al SOU de potenciar la capacidad de *autodeterminación* de los universitarios, para poder identificar, elegir y/o reconducir alternativas formativas y profesionales acordes a su potencial y trayectoria vital, a tenor de las ofrecidas por su entorno.

A ello ha contribuido en buena parte los criterios que marcan su actividad:

- Considerar de forma global e integrada necesidades y recursos, priorizando la utilización de medios disponibles.
- Anticiparse a las necesidades de sus beneficiarios directos o indirectos, acudiendo a los entornos donde se generan y sin esperar que los primeros se vean obligados a acudir al Servicio.
- Apoyar las iniciativas de todos los Centros de la División, relativas a orientación académica y profesional.
- Canalizar las demandas de orientación a través de programas de intervención, no institucionalizados a perpetuidad.

En función de los mismos, al Servicio no le ha quedado otro remedio que actuar *por programas* dirigidos a *todos* los estudiantes, en respuesta a las necesidades detectadas en sus respectivos Centros formativos y asumiendo el entorno en que éstas se producen, siendo los principios prioritarios del plan de acción los de:

- *Prevención*: Se anticipa a la aparición de circunstancias o situaciones que puedan obstaculizar el óptimo desarrollo personal, académico, social y profesional de los universitarios. Es decir, el SOU hace suyo el lema «más vale prevenir que curar», por lo que centra su actividad en las etapas previas a cada una de las fases de transición de la vida universitaria, incluida la anterior a la misma, que traspa la actuación más allá del propio entorno universitario.

- *Desarrollo*: Pretende dotar a los universitarios de las competencias necesarias, para afrontar las demandas de las sucesivas etapas de su vida, que permitan integrar cada vez más amplias y diferentes experiencias, así como proporcionar situaciones de aprendizaje vital, que faciliten la reconstrucción y progreso de sus esquemas conceptuales y actitudinales.
- *Intervención Social*: Trata de eliminar obstáculos y potenciar condiciones y oportunidades favorables para un mejor desarrollo del universitario, a través de la modificación de aspectos concretos del marco académico (políticas de admisión de estudiantes, configuraciones curriculares, metodología de estudio y trabajo, itinerarios formativos, etc.) y sociolaboral (becas y ayudas, bolsa de trabajo, itinerarios de inserción socioprofesional, etc), no tanto para ajustar o adaptar al universitario, cuanto para cambiar determinadas características del entorno, que entorpecen su desarrollo.

Por todo ello, a partir del estudio del contexto y de las necesidades detectadas en los Centros, el Servicio ofrece «a la carta» programas articulados por objetivos a lo largo de un continuum temporal. Operativamente, desde *antes* de entrar en la Universidad, al *inicio* de los estudios universitarios, *durante* y *al final* de los mismos, además del seguimiento de los itinerarios de inserción profesional, *después* de haberlos acabado.

Poco a poco, en todas ellas se busca conseguir la colaboración de servicios y entidades intra y extra universitarias, que coinciden con los objetivos prioritarios de las mismas. El logro de la misma es debido en su mayor parte a que los interlocutores perciben de que «la actuación del Servicio es necesaria y ofrece perspectivas de solución a los clásicos y actuales problemas, que más nos preocupan» (Sic).

Desde el Servicio de Gestión Académica» de la Universidad de Barcelona recibe el apoyo institucional, para el desarrollo de los programas de pre-inscripción y pre-matriculación de nuevos alumnos. El reto del futuro está en conseguir la colaboración de los programas de información y orientación, propiciados por el Departamento de Educación de la Generalitat de Catalunya en las Enseñanzas Medias.

La colaboración de los órganos de gobierno, jefaturas de estudio, profesores y estudiantes de los Centros está permitiendo optimizar los programas intersticiales. El desafío se aborda actualmente, implicando por una parte a los Centros y a sus profesores en una oferta formativa, que permita elegir con conocimiento de causa, y, por la otra, a los estudiantes «más curtidos», que puedan transmitir las experiencias vividas, para posibilitar la reducción de efectos negativos y potenciar los positivos.

Los contactos y, en mejor de los casos, los convenios con entidades directamente relacionadas con el mundo laboral (INEM, Cámaras de Comercio, empleadores habituales y nuevos de profesionales de la educación, etc.) ocupan buena parte de la actividad del SOU. Aprender a moverse en este entorno, está exigiendo al Servicio potenciar las cualificaciones necesarias, no sólo para «saber», sino además para «saber hacer» y, más aún, para «saber ser y estar».

La funcionalidad, eficiencia y eficacia de toda esta actividad, desplegada antes y después de la creación del SOU, cuenta en la actualidad con un observatorio privilegiado, como es el «Gabinete de Evaluación e Innovación Educativa Universitaria» de Barcelona, que periódicamente transmite, analiza y valora los itinerarios de inserción profesional, tras su etapa de formación.

¿TEORÍAS VS REALIDADES?

Puede decirse, que esta actividad brevemente descrita es un simple ejemplo de un servicio, pero, al fin de este suscito debate, casi es obligado recordar que, determinados ejemplos, sirven más que mil palabras.

Hoy por hoy, en España difícilmente se pueden llevar a cabo programas continuados de interven-

ción orientadora, sin contar con todas las aportaciones que pueden derivarse de los servicios, cuando se les libera de las rémoras del pasado.

Su misma existencia, puede servir a veces para contrarrestar los embates de tiempos de crisis y permitir la reconceptualización del status de la orientación, tanto a nivel institucional, como curricular.

REFERENCIAS

COMPANY, F. y ECHEVERRÍA, B. (1992): *Estructura de empleos y cualificacions en el ámbito de la orientación escolar y profesional de jóvenes y adultos en España*. Luxemburgo: Oficina de Publicaciones Oficiales de las Comunidades Europeas. (CEDEFOP).

ORIENTACIÓN POR PROGRAMAS: UN CASO PRÁCTICO

Josefina Hernández Fernández
Universidad de Murcia

Me corresponde en esta ocasión apoyar con un caso práctico las ventajas que la intervención por programas tiene para el logro de los fines, objetivos y cumplimiento de funciones de la Orientación Educativa. Pero antes de pasar a exponer el programa, y partiendo de la opinión de que ambos modelos no son incompatibles (al igual que ocurre con la mayoría de los modelos en Orientación), quiero resaltar, aunque brevemente, las ventajas que el modelo de programas mantiene frente al de servicios y por ende resaltar las cualidades y características que, en este sentido, tiene el programa que presentamos¹.

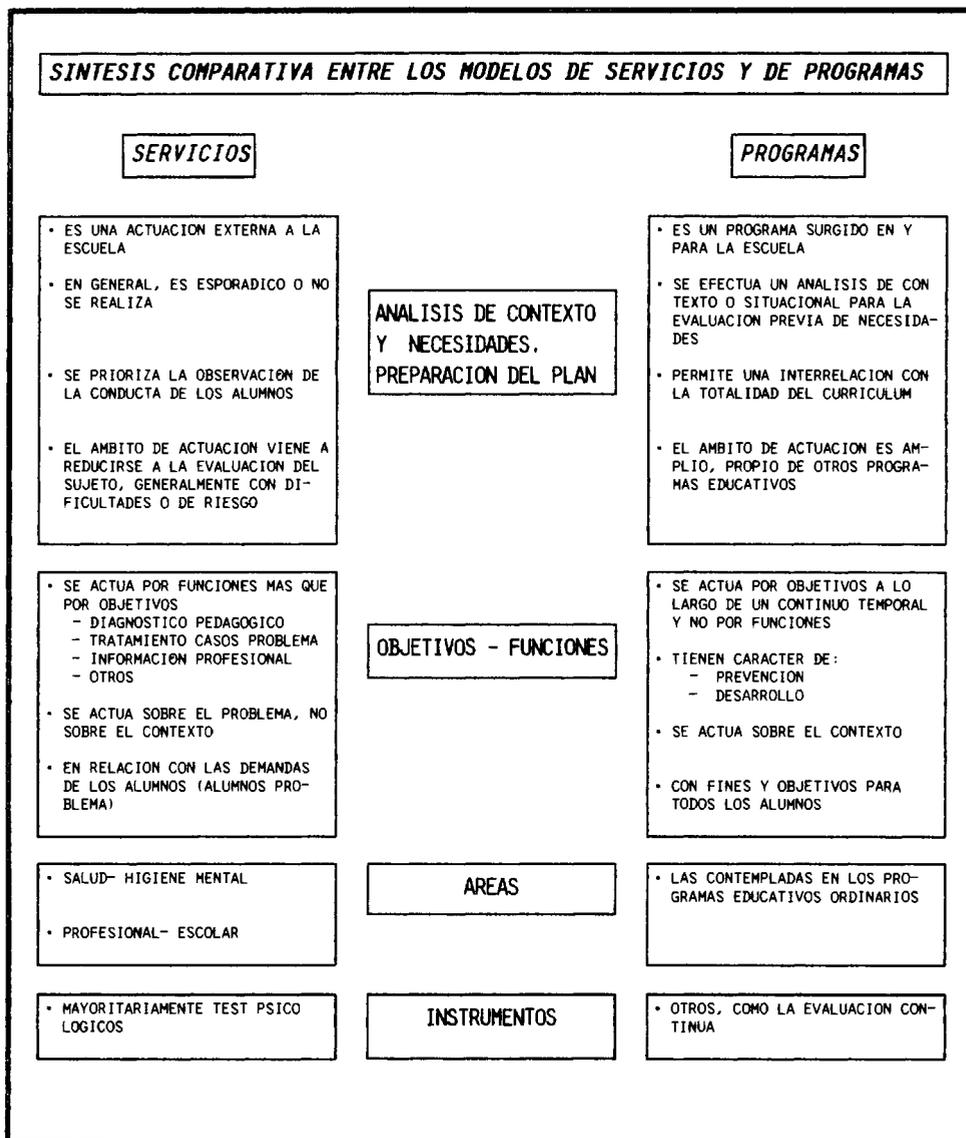
En el cuadro que se incluye a continuación se pueden ver contrastados algunos aspectos de ambos modelos, de cuya comparación se puede extraer algunas de las posibles ventajas y limitaciones de ambos modelos para su aplicación en Centros Educativos y sobre todo, como indicábamos antes, para la consecución de los objetivos orientadores.

Por señalar algunos de ellos, el enfoque clínico, reactivo y basado en las necesidades del sujeto (y sobre todo en lo que ha sido su especificación española en los sujetos con necesidades especiales), dificulta cuando no impide mantener la suficiente atención a los aspectos preventivos de la intervención orientadora, por no hacer referencia al clásico principio de la orientación para todos los alumnos; el apoyarse en el equipo multidisciplinar y en las competencias del técnico y no en el equipo surgido en el propio centro, integrando a profesores y agentes implicados en la tarea orientadora, refuerza la marginación de la Orientación del proyecto educativo del profesor y de ahí, se hace más difícil la deseada y necesaria implicación de los profesores en dichas tareas.

Otros comentarios se podrían hacer de las distintas limitaciones del modelo de Servicios a tenor de los aspectos recogidos y sobre todo en lo que ha sido su puesta en marcha en nuestro contexto y que han sido suficientemente remarcadas por otros autores: Álvarez, 1988; Rodríguez Espinar, 1986; Sanz Oro, 1990).

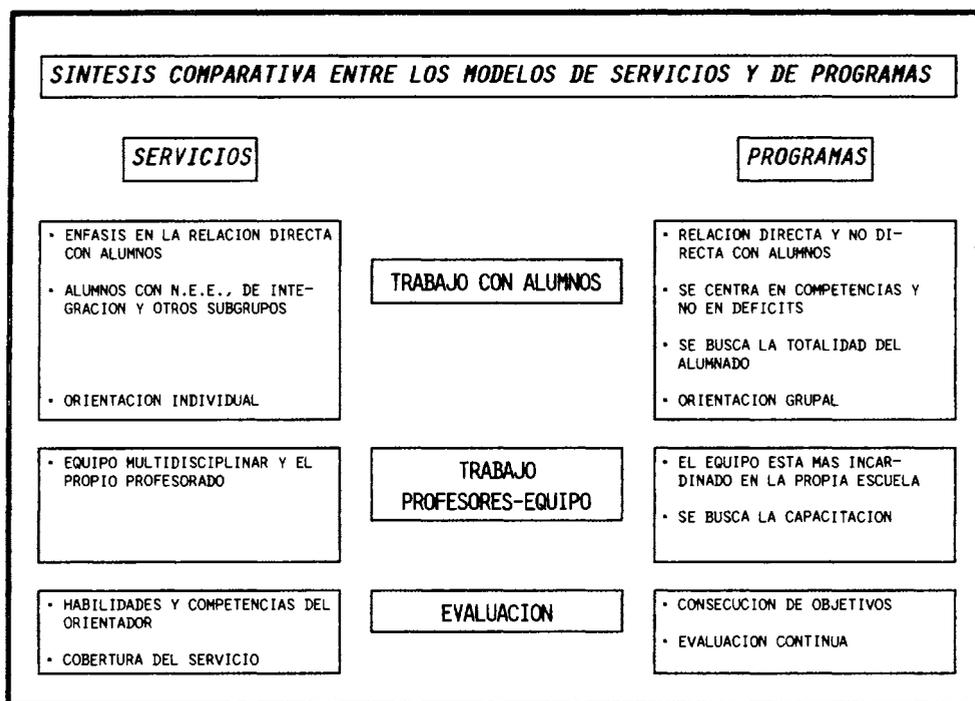
Sin embargo, desde la óptica práctica en que se sitúa este comentario, aunque el modelo de Servicios encierra serias limitaciones para el logro de los objetivos orientadores, no se puede negar que el modelo por Programas, aunque con claras ventajas en este sentido (centrarse en las necesidades del

¹ En este caso el plural no es meramente mayestático o estilístico. El programa ha sido diseñado conjuntamente por Pilar Martínez Clares y yo misma y experimentado y puesto en marcha por ella como base de una investigación vías de realización.



centro y del aula-clase, implicar a los agentes educativos, permitir más fácilmente la prevención y la atención a los aspectos de desarrollo del sujeto...), también encierra todavía ciertas limitaciones para su aplicación en nuestro contexto, si lo que se pretende es su puesta en marcha ateniéndose a la totalidad de sus premisas teóricas.

Sobre todo en lo que se refiere a sus implicaciones curriculares y la preparación y capacitación de los agentes orientadores. De ahí que en la puesta en marcha del programa que presentamos, se haya obviado su inclusión en los currícula de los alumnos y se haya optado por plantearlo como un



programa de acción tutorial: dirigido a tutores y desarrollado por ellos, previa la capacitación de los mismos. Veamos algunas de sus características.

UN PROGRAMA DE DESARROLLO PERSONAL Y SOCIAL: EL AUTOCONCEPTO

La elección del tópico del programa viene apoyada por las actuales consideraciones sobre la intervención psicopedagógica en el nuevo tramo educativo que comprende la Educación Secundaria Obligatoria, que suponen una renovación para dicha intervención en lo que se refiere a Programas de Desarrollo Personal y Social, en su vertiente de conocimiento de sí mismo (self-concept) para alcanzar la madurez personal y el equilibrio afectivo, asignándole cada vez más unánimemente al autoconcepto una función importante en la integración de la persona, la motivación del comportamiento y el desarrollo de la salud mental.

El programa que presentamos está basado en el Programa PATSS, un programa de Actitudes hacia la escuela y hacia sí mismo, apoyado en las propuestas del «*Building a Positive self-concept*» de Jacobs, Turk y Horn.

El programa esta destinado a alumnos/as que se encuentran en la Enseñanza Secundaria Obligatoria; estos preadolescentes y adolescentes viven en esta etapa un proceso de formación de una identidad donde confluyen situaciones diferentes y conflictivas.

Está estructurado para que pueda ser el profesor/a tutor/a quien lo lleve a la práctica, recomendándose que sea en colaboración y asesoramiento con el orientador/a del centro, ambos en estrecha cooperación deberán diseñar, planificar y evaluar las distintas sesiones de trabajo, que requieren por su parte un horario estable dentro del calendario escolar, sugiriéndose la hora destinada a tutoría.

El objetivo general o finalidad del programa es el construir o desarrollar un autoconcepto positivo en los estudiantes. Las actividades o estrategias están estructuradas para que los estudiantes tomen conciencia de la relación existente entre cómo se perciben a ellos mismos y cómo se comportan. Se pretende que los alumnos/as se responsabilicen de sus elecciones y practiquen técnicas que fomenten su autoestima y autoconcepto.

Si bien los autores del PATSS recomiendan utilizar el programa diariamente al comienzo del año escolar para desarrollar la confianza y sentido de comunidad en la clase, ya que esto contribuye a desarrollar un aprendizaje positivo del entorno y utilizado tempranamente ofrece información sobre cómo los estudiantes sienten sobre ellos mismos, y por tanto proporciona una oportunidad para hacer frente a las necesidades individuales y ayudará también a decidir el mejor orden en el que usar los ejercicios para lograr una mayor efectividad y cohesión en el grupo, las estrategias y actividades del programa que presentamos pueden ser utilizadas con efectividad en los espacios periódicos dedicados a la Tutoría.

ESTRUCTURA DEL PROGRAMA

El programa se divide en **cuatro apartados o núcleos temáticos**:

I. Aprendiendo a Conocer a los Otros, consiste en una serie de actividades que están previstas para crear un clima confianza, empatía y un sentimiento de comunidad entre los estudiantes de la clase. Además proporciona oportunidades para desarrollar las habilidades de comunicación verbal y no verbal.

Los principales objetivos de aprendizaje de este bloque son: a) Ayudar a aprender a conocerse y a sentirse mejor con ellos mismos. b) Ayudar a conocer a los otros y a desarrollar la confianza entre compañeros/as. c) Ayudar a prestar atención a los sentimientos.

II. Aprendiendo a Conocerte, presenta estrategias que están específicamente diseñadas para ayudar a desarrollar la autoconciencia al reflexionar sobre las emociones, sentimientos y valores.

Los principales objetivos de aprendizaje de es bloque son los que siguen:

a) Ayudar a los alumnos/as a identificar y aceptar sus sentimientos y animar a expresar sus sentimientos a los iguales.

b) Incrementar la conciencia del autoconcepto en los alumnos/as al verse a si mismo como los otros los ven.

c) Ayudar a ver los beneficios de las situaciones difíciles o decepcionantes y a desarrollar actitudes positivas.

d) Ayudar a comprender los diversos papeles que pueden interpretar con diferentes personas y en diferentes lugares.

e) Ayudar a relacionar el autoconcepto con los valores y a identificar los rasgos positivos y negativos.

III. Aprendiendo a Parecerte a ti mismo, se centra en actividades que aumentan el autoconcepto al desarrollar actitudes positivas. A través de autotareas, visualización y autoafirmación, se pretende enseñar técnicas de autocontrol tales como el dominio del estrés.

Aquí los objetivos de aprendizaje destacados son:

a) Ayudar a los estudiantes a ver la relación entre auto-conservación (hablar con uno mismo), emociones y conducta y expresar los sentimientos que resulten.

b) Conocer cómo influyen los otros en nuestro autoconcepto.

PROGRAMA DE AUTOCONCEPTO	
SESION NUM.:	
BLOQUE I	
APRENDIENDO A CONOCER A LOS OTROS	
FINALIDAD	
OBJETIVOS DEL APRENDIZAJE	
ESTRATEGIAS Y ACTIVIDADES	
HOJA DE TRABAJO	
<p>Crear un clima de confianza, empatía y un sentimiento de comunidad entre los alumnos/as de clase, proporcionando oportunidades para desarrollar habilidades de comunicación verbal y no verbal.</p>	
<p>a) Ayudar a aprender a conocerse y a sentirse mejor con ellos mismos. b) Ayudar a conocer a los otros y a desarrollar la confianza entre compañeros/as. c) Ayudar a prestar atención a los sentimientos.</p>	
<p>1. ¿Quién eres tú?</p> <p>2. Cuestionario autobiográfico.</p> <p>3. Círculos dobles.</p> <p>4. Constelación de símbolos.</p> <p>5. Observaciones.</p> <p>6. Compartir los propios sentimientos.</p> <p>7. Auto - Bio - Rompehielos.</p> <p>8. La entrevista.</p> <p>9. ¿Cuál es mi especialidad?.</p>	<p>1</p> <p>1</p> <p>2</p> <p>3</p> <p>3</p> <p>4</p> <p>4</p> <p>5</p> <p>5</p>

- c) Ayudar a identificar medios para tratar situaciones estresantes.
- d) Animar a desarrollar planes personales para hacer frente al estrés en sus vidas.

IV. Listos para el Futuro, se dirige a las habilidades para la vida: establecimiento de metas, toma de decisión, solución de problemas, resolución de conflictos, manejo de la presión de los iguales, tratos con la autoridad, liderazgo y toma de responsabilidad.

Los principales objetivos de aprendizaje son: a) Ayudar a desarrollar una definición del concepto de meta y comprender la diferencia entre meta a corto y largo plazo. b) Aprender un proceso de cuatro pasos para hacer frente a los conflictos y aprender a solucionar los problemas.

Estos bloques de contenidos van precedidos de un bloque preliminar de análisis y exploración del nivel del autoconcepto y autoestima de los sujetos participantes en el programa y seguidos de un bloque final de exploración y valoración para el análisis y valoración del autoconcepto y autoestima así como para la evaluación del programa.

Cada bloque va acompañado de una hoja-guía para el tutor en la que se sugiere las posibles técnicas a utilizar para la consecución de los objetivos (incluimos una de ellas a modo de ejemplo). A la vez que se facilita un cuaderno de trabajo para los alumnos con las distintas técnicas y actividades a realizar.

El programa se desarrolla en un número indeterminado de sesiones, según las necesidades y el tiempo que se disponga si bien se sugiere una posible temporalización del programa lo largo del curso. Se poseen un gran variedad de actividades o estrategias para poder ir trabajando un bloque determinado hasta que se consigan o afiancen los objetivos de aprendizaje, aunque se pueden trabajar simultáneamente distintos bloques temáticos. Así mismo, se facilitan materiales para la evaluación de las sesiones de trabajo y el logro de resultados.

CONSIDERACIONES FINALES

La utilización del modelo de programas ha permitido entre otras cosas y a tenor de lo que estamos defendiendo en esta intervención:

- La participación del centro educativo, desde el momento en que el programa fue presentado al Consejo Escolar, aprobada la intervención y su inclusión en el proyecto educativo del centro a partir de considerarlo un programa del Departamento de Orientación.

- Partir de las necesidades del Centro, pues la intervención fue demandada como consecuencia del análisis de necesidades, surgido del análisis de contexto previo realizado por el Departamento de Orientación del Centro.

- Realizar una intervención de claro matiz preventivo, basada en el desarrollo de los alumnos, dirigida a todos y no sólo a los que presentaban déficits, pero que a la vez ha tenido en cuenta el punto de partida de los sujetos a los que iba dirigida.

- Facilitar la capacitación de los tutores, a través de asumir el orientador el papel de consultor/educador de los mismos.

- Posibilitar el trabajo en equipo (del que el orientador ha sido considerado un miembro más), puesto que el programa se ha desarrollado en todos los cursos del ciclo superior del centro y el intercambio de impresiones y puestas en común han sido periódicas.

- Ha posibilitado la realización de una acción intencional, secuenciada, prolongada a lo largo del curso y con la evaluación tanto del proceso como de los resultados

- Ha permitido la colaboración de otros profesionales en la puesta en marcha del programa

Difícilmente se habrían podido cubrir estos aspectos desde el esquema de actuación del modelo de Servicios.

SIMPOSIUM II

¿POR QUÉ DIFERENCIAR LOS TÉRMINOS EVALUACIÓN Y DIAGNÓSTICO?

Coordinador:

P. Municio

(Universidad Complutense de Madrid)

Participantes:

I. Alfaro Rocher

(Universidad de Valencia)

E. García Jiménez

(Universidad de Sevilla)

R. Pérez Juste

(Universidad Nacional de Educación a Distancia)

A. Lázaro Martínez

(Universidad Complutense de Madrid)

DIAGNÓSTICO Y EVALUACIÓN

por

Ignacio J. Alfaro Rocher

Universitat de València

Básicamente partiré de lo esencial del concepto de diagnóstico y de lo esencial del concepto de evaluación. Sin embargo, no pretendo realizar afirmaciones que sean válidas para todos, al contrario, quiero presentar desde mi experiencia y ensayos de investigación lo que entiendo por cada concepto para mi trabajo. Si mi enfoque sirve para otros podrá considerarse como una aportación más a estos conceptos.

COMPRESIÓN Y EXTENSIÓN DE LOS CONCEPTOS

A veces nos preguntamos, qué es más comprensivo, el diagnóstico o la evaluación. ¿En relación a qué objeto u objetivos radica esa comprensibilidad?, ¿Que concepto incluye más?, ¿Cuál de ellos resalta más los aspectos positivos de los sujetos o situaciones? etc. Para evitar preguntas tan disparatadas habría que definir previamente lo esencial de los conceptos. Los conceptos tienen una comprensión y una extensión. La comprensión son las notas o características que definen un concepto y la extensión es a qué número de sujetos o de clases se extiende o abarca el concepto. Por ejemplo, *hombre=animal racional* es la comprensión y, a qué seres se aplica, es la extensión. La comprensión debería fijarse en una definición esencial, mediante notas que lo caractericen y lo distingan o separen de otro concepto. Sin embargo, cada nota que se le añade a la comprensión de un concepto hace que éste abarque menos, que se vaya restringiendo. Leibniz distingue dos tipos, internamente, las que expresan la realidad de ese ser y, externamente, las que lo distinguen de los otros. La primera es claridad interna del concepto, poner notas que claramente digan lo que es. Es lo que hay que hacer con nuestros conceptos: distinguirlos radicalmente. La definición no dependerá ni del para qué, ni será más o menos comprensible en función de nada.

Según mi criterio, evaluar es asignar un valor a una cosa objetivamente dada y de conformidad con un criterio. No incluyo nada que haga referencia a lo cualitativo o cuantitativo para que el concepto tenga toda su riqueza, sin restricciones. Hay situaciones que no son cuantificables y por ello no dejan de ser evaluables, sin embargo, la evaluación es valoración, ponderación (*pondus*). Hablo de evaluación cuando se busca el *cuánto*, si mucho o poco, en relación a un criterio suficiente, oportuno, adecuado, cualitativa o cuantitativamente en relación con un tipo de valor u otro. En síntesis, se trata de la *búsqueda del cuánto* en relación con un módulo valorativo.

Por el contrario, entiendo por diagnóstico la *búsqueda del qué*. La búsqueda de cuales son las causas, síntomas, indicios, etc. de tanto un ajuste como un desajuste. Es decir, no tiene por qué tratarse de una situación patológica, aunque generalmente se suele pedir un diagnóstico cuando la situación es problemática. Cuando un niño está correctamente encaminado, nadie busca el diagnóstico, lo que se hace es evaluar, dejar constancia de ese comportamiento.

Creo que esta primera distinción sintética, que a veces puede resultar parcialmente falsa, es fundamental antes de aplicar estos términos a la educación, que sería la extensión de los conceptos. Téngase en cuenta que ambos términos son analógicos (medicina, economía, etc.), no exclusivos de ningún área de conocimiento, y se hace necesario precisar lo que se entiende por ellos para evitar la usual contaminación con otras áreas no educativas.

LA BÚSQUEDA DEL QUÉ Y EL CUANTO PEDAGÓGICOS

A ambos conceptos se les añade el adjetivo «pedagógico» y, a veces, los consideran como tal por el simple hecho de que se dan en la escuela o el objeto son los escolares, aunque las descripciones sean psicológicas o sociológicas. Incluso, los más estrictos, se sitúan en el polo opuesto y hablan de su esencia «pedagógica» porque el método utilizado se aparta de otros, de la exploración psicométrica clásica, del análisis conductual, etc. No opino lo mismo. Es la *intención* la que hace que se elija de todo lo que describe un fenómeno lo que es útil para modificarlo, lo que es educativo, lo que procura el desarrollo y mejora del niño. Si se aplica el criterio pedagógico al diagnóstico, es porque se ha seleccionado de un conjunto fenomenológico de variables, cualesquiera que sean, aquellas sobre las cuales influir para modificar un comportamiento o un procedimiento de enseñanza que subsane los fallos, los errores, etc. Lo pedagógico no es el método utilizado o la situación escolar sino la intención, la acción meliorativa, el acto. Si se aplica el adjetivo a la evaluación, es porque está encaminada a

mejorar la enseñanza, a resaltar qué método o instrumento es más eficaz y porque sus valoraciones reflejan mejor el valor educativo o se acomodan más a los parámetros educativos que pretendemos. No entiendo por qué algunos hablan de que un concepto está más integrado o cercano que el otro al proceso didáctico, al contexto del aula, al profesor, al especialista, a la orientación o a la didáctica. Al diagnóstico, al menos, no se le puede limitar ningún campo o área temática que pueda ser un síntoma, un síndrome (σύνδρομοζ= que corre a la par, que concurre). Cualquier cosa puede ser un indicio.

Lo importante es la intención didáctica de ambos conceptos, tanto uno como otro están encaminados al «διδασκαλοζ», a la habilidad del enseñante, lo que significa que ambos pueden analizar cualquier proceso, institución, sujeto o agente educativo desde un punto de vista diferenciado, pero con una finalidad de mejora. Existen métodos que proponen influir de una manera o influir de otra, la evaluación intentará determinar cual es el más eficaz, el más pedagógico. El diagnóstico, por su parte, será pedagógico por cuanto su análisis de las causas está encaminado a influir en aquellos elementos esenciales o por cuanto propone unos principios pedagógicos, un método de intervención, que procura un desarrollo deseado.

ANTERIORIDAD Y POSTERIORIDAD DE LOS EFECTOS

Existen aspectos diferenciales entre la evaluación y el diagnóstico. Decíamos al principio que en la evaluación, en la búsqueda del «cuánto», ese algo le era «dado» objetivamente al que valora. En cambio en el diagnóstico, ese algo es un «qué» que tiene que descubrir, que debe modificar. No se le da hecho. Lo que encuentra son síntomas que pueden ser simultáneos, estar todos presentes en un momento dado (διὰ= a través de, totalmente) o estudiarse en un proceso (διὰ= a lo largo de). Tiene que recorrer cada proceso de enseñanza en sus características básicas. Por ejemplo, en el proceso de aprender a escribir, que tiene sus pasos, unas letras tienen su curso, unas estructuras figurales se suceden a otras, etc. la evaluación nos dirá constantemente en qué momento se sitúa un individuo dentro de ese proceso, si su desarrollo alcanza lo esperado, en cambio el diagnóstico, intentará descubrir cual es la naturaleza del e'r cometido escribiendo y propondrá un camino educativo para continuar con el progreso adecuado de la escritura en el niño.

En otras palabras, mientras que en el diagnóstico el efecto del proceso es posterior al análisis, se interviene para mejorarlo o cambiarlo, en la evaluación el efecto es anterior a la evaluación. La evaluación para verificarse necesita tener presente lo que se está evaluando en su ser realizado. Se evalúa lo que ya existe. Incluso desde un enfoque cualitativo, los efectos se van produciendo o están presentes a la vez que se interviene en un proceso. La intervención será dinámica y producirá unos efectos, pero son los efectos lo que se evalúa y, en todo caso, si la intervención es eficaz en relación a algún criterio. Sin embargo, eso no es averiguar las causas del error, o averiguar las causas de los efectos que se van produciendo. En el diagnóstico se interviene en lo que queremos que se produzca, sobre procesos en marcha por supuesto, pero con una intención proactiva. Así, la evaluación actúa sobre realidades, productos o procesos, y el diagnóstico sobre los sujetos o contextos para que se produzcan los efectos deseados.

DINAMISMO PROCESUAL E INFLUJO ENTRE LOS DOS TIPOS DE BÚSQUEDA

Para algunos, puede resultar incómodo el término «dado» y más si ello puede poner en peligro el carácter procesual y continuo. Sin embargo, no se deduce tal cosa. Entiendo que el proceso de evaluación empieza cuando proceso de elaboración si no, no se evalúa nada. En último extremo, tanto desde un paradigma cualitativo como cuantitativo de la evaluación existe un dinamismo evaluador, al niño se le van dando unas pautas dentro de un proceso y, en función de los resultados, se cambia o no

la estrategia de intervención. No hay que establecer una zanja en el proceso evaluativo sino en los puntos de vista de abordar el proceso.

Esto significa que la evaluación, aun siendo continua, se va evaluando efecto tras efecto, y así se puede seguir el curso del individuo, del proceso, etc., pero eso no es lo mismo que averiguar por qué fenómeno, por qué razón se están produciendo unos efectos y no otros. ¿Qué le está pasando a un niño cuando intenta copiar la «K» y su resultado se asemeja a una «F» o son dos paralelas verticales? Eso no es evaluar sino averiguar la razón de ese resultado, el niño está afectado por algo que le produce esos efectos, que son los indicios, antecedentes, causas que producen unos efectos no deseados. Consecuentemente, se diseña una estrategia de intervención como efecto de un diagnóstico y se continúa sondeando el resultado mediante la evaluación. Una cosa es que el diagnóstico pueda ser evaluable y otra cosa que el diagnóstico consista en la evaluación. Un ciclista puede ser bueno o malo, pero ser ciclista no consiste en la evaluación que se haga de él sino en pedalear correctamente y si se tiene algún defecto descubrirlo, diseñar una estrategia apropiada, etc. Para ello, es necesario un diagnóstico mientras pedalea, hay que hacerle pedalear y modificar aquello que no realiza bien. En ese momento no se está evaluando, se está diagnosticando.

Otra cuestión es el influjo conjunto, el efecto conjunto entre el diagnóstico, que determina lo que hay que hacer y la evaluación que pondera si se produce el efecto deseado o suficiente, o también si en vez de una intervención, previo otros diagnósticos y evaluaciones, conviene diseñar otras. Es decir, se puede evaluar el efecto, si la manera de hacer el diagnóstico es adecuada o no, pero siempre ocurrirá sobre algo que le es «dado» al evaluador. En cambio al diagnóstico le es dado el fallo, la necesidad de hacer un estudio, un análisis. En cualquier caso, hay un influjo conjunto y dinámico entre las dos formas diferenciadas de análisis. La evaluación siempre actúa sobre lo encontrado, no encuentra; sin embargo, si como consecuencia del diagnóstico lo que se busca es tender a producir un efecto, habrá que valorar si se produce o no. La evaluación pondera los efectos educativos, los fenómenos, que tienen en su principio un diagnóstico, una razón que obliga a modificar el sistema. El diagnóstico y la evaluación se llaman mutuamente a través de un sistema dinámico o *procedimiento de ajuste* en busca de los efectos deseados. El método de ajuste psicológico consiste en eso, en subir y bajar los parámetros hasta encontrar el método adecuado de medida e intervención. Al mismo tiempo que el diagnóstico descubre los aspectos que hay que abordar para remediar un efecto o decide qué instrumento y sistema aplicar, la evaluación debe controlar el efecto.

Este principio de ajuste, aplicado a nuestros trabajos sobre escritura y a nuestro método (Secadas, Rodríguez y Alfaro, 1993) queda suficientemente claro. Contamos con escalas, baremos, medidas de velocidad, índices modulares de las escalas, etc. que permiten al profesor, en cualquier parte del proceso, situar a los sujetos. Todo ello sirve para evaluar el efecto de nuestro método de enseñanza de la escritura, son procedimientos de valoración. Existe un estándar, en unos casos valorado y en otros estimado por el profesor, sobre el que se determina o evalúa la bondad del método. Sin embargo, el hecho de decir a qué se debe que un niño esté haciendo mal una letra o grupo de letras, las fases por las que transcurren los errores, su causa, el proceso para reconducir la intervención, etc. sería el diagnóstico en sus dos acepciones del prefijo «διδά»: a) en un momento dado, a través de instrumentos estáticos (teniendo presentes todos los indicios, los síntomas) y b) procesual, a través de una secuencia que busca una estructura unitaria de interpretación siguiendo un proceso. En cualquier caso, todo lo psicológico y pedagógico es procesual aunque se establezca en determinados momentos en los que se observa y se busca la pista dentro de un proceso. Eso es el síndrome, las cosas que van coincidiendo, que se mantienen «junto a», o que acompañan al error (sundromh, concomitantes al curso del fenómeno, que concurren). De tal análisis de errores se extraen las causas e interviniendo sobre el proceso, en uno u otros elementos, se podrá conseguir un desarrollo más óptimo del sujeto en escritura. Asimismo,

la evaluación va controlando si se consigue o no, aunque siempre tiene que tener el efecto previamente registrado. En tal sentido es continua porque recorre las partes del proceso.

CONOCER Y DESCRIBIR

Parece redundante plantearse estas cuestiones siendo que nos referimos a dos ámbitos de investigación. Sin embargo, existen matizaciones y embrollos que prefiero aclarar desde mi punto de vista. La evaluación pretende conocer la aceptabilidad de un sistema en relación a criterios, describir sus elementos, estructura, fenómenos educativos, establecer puntos de consecución de dominios, etc. y tomar decisiones de mejora, educativas. Del mismo modo, el diagnóstico necesita de ese conocimiento, y de otros que lo traspasan, para determinar las causas (internas o externas a los sujetos) de los efectos deseados o no y, consecuentemente, tomar decisiones educativas. La clave diferencial del objeto de conocimiento sigue siendo la misma que se definió al principio, el *cuánto* y el *qué*. Por el contrario, a veces, se suele adscribir a la evaluación el conocimiento de las causas, las decisiones diagnósticas, las correctivas, etc. o al diagnóstico la actividad de conocer y describir los efectos o de «*conocer las causas subyacentes o adscribir a cuadros nosológicos conocidos*». Las primeras cuestiones creo que están resueltas en apartados anteriores, sin embargo, voy a referirme ahora al último entrecomillado ya que su resolución no depende del discurso anterior sino de mi personal enfoque teórico.

En diagnóstico, estamos acostumbrados a convivir con un panorama consensuado que traslada al ámbito educativo los enfoques médicos y psiconeurológicos clásicos patologizando las dificultades de aprendizaje, buscando explicaciones subyacentes o intentando una adscripción a cuadros nosológicos conocidos. El problema está precisamente ahí: «en tratar a los sujetos pertenecientes a un grupo determinado como si fuesen relativamente homogéneos». Existe poca información que pueda ser empleada para «clasificarlos» de un modo preciso, más aún, los patrones son múltiples y distintos, no aparecen caracterizaciones precisas (Campione y otros, 1982). En psicología, los prolijos cuadros nosológicos, además de difusos, presentan multitud de causas. La descripción diagnóstica no depende de una adscripción determinada sino del *qué*, de la búsqueda de las causas en un momento concreto y ante un hecho determinado. En relación a las causas subyacentes, es también usual la exploración de algunas áreas como son, la personalidad, los aspectos psicológicos, emocionales, biológicos, los factores situacionales y, como último reducto, el rendimiento académico. Todo ello en busca de factores subyacentes o áreas más o menos deficitarias que pudieran dar alguna explicación. ¿El resultado? No es el momento en este debate de reproducir ahora la cantidad de trabajos científicos que han evidenciado el fracaso de estas aproximaciones. Mi opción personal, por el contrario, si bien no descarta ninguna posibilidad que pueda explicar algo, tampoco da prioridad al conocimiento y a la descripción de los factores subyacentes.

Naturalmente si las causas están en un individuo, que no lo están siempre, no pueden dejar de ser psicológicas, pero todo lo que está en un individuo, aunque sea psicológico, no lo es totalmente. El comportamiento de un hombre puede ser moral o de otra índole, pero en cuanto moral, se sale de la psicología, es de otro orden. Sin embargo, lo pedagógico se monta sobre la dinámica de un sujeto, sobre lo psicológico. Es decir, para hacer un diagnóstico no se puede excluir lo psicológico, pero no se estudia, no se analiza para conocer los procesos psicológicos, sino para conocer en qué medida los efectos pedagógicos dependen de los psicológicos, que es distinto. La psicología de la educación no estudia los fenómenos de forma neutra u objetiva, para luego aplicarlo a la educación, sino que de toda la fenomenología psicológica ve qué aspectos, fenómenos o procesos son susceptibles de intervención o de modificación para mejorar el aprendizaje o el estado del sujeto. Los que influyen, los que son eficaces, pueden constituir un corpus doctrinal aparte, científico, que refleja las normas que dirigen o

regulan esa eficacia. La regulación de esa eficacia, la normación de esa eficacia es lo psicopedagógico, pero no lo psicológico en sí. Yo entiendo que si un niño tiene un problema, por ejemplo, dificultad en escritura y se empieza a estudiar y conocer la percepción, la visomotricidad, la psicomotricidad, la memoria, etc. y con todo ese perfil, apreciado por tests, mediante una ecuación de regresión múltiple, se puede predecir el comportamiento del niño, no se ha hecho un *diagnóstico pedagógico* sino una *descripción psicológica*, ya que a base de factores puramente psicológicos, es probable predecir el comportamiento, lo que no quiere decir que interviniendo en el test de memoria, por ejemplo, pueda obtenerse nada. Recuérdese la inconsistente historia de los entrenamientos psiconeurológicos o los problemas de transferencia de aprendizajes. En definitiva y desde mi punto de vista, hay que diagnosticar en base a las cosas que son investigables dentro del camino de la pedagogía, aunque tengan un soporte psicológico por que están en el sujeto.

LA METODOLOGÍA DEL ERROR EN EL DIAGNÓSTICO PEDAGÓGICO

El diagnóstico permite ver las variables sobre las que hay que actuar o que denuncian el tipo de error que se está cometiendo. En otro sentido, el análisis del tipo de error puede guiar la intervención educativa: ¿se trata de un retraso específico, de una mala estrategia, de un retroceso a automatismos más simples o simplemente, de optimizar un comportamiento que es mejorable? El error es la pista por la que se desvía el niño, hay que averiguar por qué se deriva, cuales son los indicios que conducen al error. Todo ello justifica el diagnóstico pedagógico. El diagnóstico basado en el análisis de errores, sin obviar otras situaciones de enseñanza-aprendizaje u otras variables psicológicas, biológicas o sociales, repercutirá en el diseño de una intervención educativa.

El análisis de errores es un instrumento metodológico utilizado por la Psicología cognitiva en el estudio de los procesos de información. Así, encontramos autores que conceden un gran valor a los errores en el aprendizaje, ya que reflejan el procesamiento que opera en el sujeto: «La inteligencia consiste en comprender cosas desde el punto de vista de lo que uno ya sabe, con la memoria impulsada por el error (*Failure Driven Memory*) como fuente de crecimiento de la memoria...» (Dehn y Shank, 1982). Siegler y Richards (1982), por otra parte, en su estudio evolutivo de la inteligencia extraen ciertas conclusiones que conducen a una forma de ver la inteligencia como desarrollo de habilidades relacionadas con el aprendizaje. A los autores les sugiere un tipo de planteamiento, el *rule-assesment* que pudiéramos traducir por adiestramiento en la regla o, de modo más explícito, *adquisición progresiva de la norma correcta*. Su base teórica puede resumirse en diversos puntos, entre los que destacamos que, la eficacia del aprendizaje depende, en gran parte, de que el niño distinga la diferencia entre lo que hace y la norma correcta, entre el error y el acierto. La *metodología del error*, como la llaman los autores, «si bien ha sido aplicada, no ha sido sometida a riguroso estudio, al menos en Psicología cognitiva, debido sin duda a la postergación sufrida por el estudio del proceso de aprendizaje, englobado injustamente entre las tareas propias de un enfoque conductista». Nosotros en escritura hemos utilizado el análisis de errores (Secadas, Rodríguez y Alfaro, 1993), no sólo en su versión *cognitiva* pura —registrando los fallos cometidos en la percepción y ejecución de las figuras— sino en la dimensión evolutiva, indagando cómo progresa la elaboración de las letras, de qué modo se va perfeccionando su ejecución, etc. Así, podemos encontrar errores progresivos, que marcan un avance en la adquisición de una habilidad, y errores regresivos, en el sentido de retroceso ante el fracaso. En escritura hemos diferenciado cuatro tipos de errores muy comunes que denominamos: error por retraso, error regresivo o por fracaso, error estereotípico y error de acabado. Este análisis no es evaluativo sino diagnóstico: se trata de detectar el tipo de error, localizarlo en una fase o secuencia evolutiva y recomendar el tipo de actuación educativa más apropiada. La evaluación determinará si se consigue el efecto deseado o si es preciso un nuevo ajuste dinámico.

BIBLIOGRAFÍA

- CAMPIONE, J. C.; BROWN, A. L. y FERRARA, R. A. (1982): «Retraso mental e inteligencia» en R. J. STERNBERG (1982) *Inteligencia humana, II. Sociedad, cultura e inteligencia*. Buenos Aires: Paidós
- DEHN, N. y SCHANK, R. (1982): «Inteligencia humana e inteligencia artificial» en R. J. STERNBERG (1982): *Inteligencia humana, II. Sociedad, cultura e inteligencia*. Buenos Aires: Paidós.
- SECADAS, F., RODRÍGUEZ, M. T. y ALFARO, I. (1993): *Escribir es fácil. Bases psicológicas y experimentales de una nueva metodología*. Madrid: Ed. TEA S.A., Depósito Legal de Madrid 9128 y 9129. (2 vols.). En prensa.
- SEGLER, R. S. y RICHARDS, D. D. (1982): «The development of intelligence», en R. J. STERNBERG (1982): *Handbook of Human Intelligence*. New York: Cambridge University Press.

¿DIAGNÓSTICO VERSUS EVALUACIÓN?

por
Eduardo García Jiménez
(Universidad de Sevilla)

En este debate en torno al significado de los conceptos de diagnóstico y evaluación, nos gustaría construir nuestra argumentación teniendo presente tres supuestos. El primero de ellos, señalaría que el significado que se atribuye a los conceptos —diagnóstico y evaluación— está condicionado por el modus operandi asociado a los mismos y por el contexto concreto en que se aplican.

Los conceptos, en el marco de las ciencias sociales, son acción y práctica humana mediatizadas por un contexto. Acercarse a ellos supone hacerlo al modo en que se utilizan aquí y allá. Tratar de delimitar las acciones y prácticas propias de la evaluación y del diagnóstico en educación requiere, por tanto, considerar ciertos elementos personales y contextuales de los que va a depender el significado de esos conceptos: (a) el marco educativo, legal e institucional (incluso el ámbito público o privado) en el que se va a desarrollar la evaluación o el diagnóstico; (b) la perspectiva teórica (formación inicial y conocimiento práctico desarrollado) del profesional que realiza el diagnóstico y/o la evaluación; (c) la cultura asociada a ambas prácticas, etc. De qué otro modo podrían explicarse nuestra dificultad para entender las diferencias entre las prácticas asociadas a conceptos (*assessment*, *evaluation*, *appraisal*, *diagnosis*, *diagnostic assessment*, *evaluation research*, *accountability*, *accreditation*, etc.) utilizados por profesionales que trabajan en otros contextos educativos, con una formación, una cultura y unas experiencias diferentes; y, de qué otro modo podría explicarse nuestra dificultad para entendernos con otros profesionales que utilizan los conceptos de diagnóstico y evaluación fuera del ámbito universitario.

Un segundo supuesto, del que partimos al definir los conceptos de evaluación y diagnóstico en educación, es que ambos conceptos sólo tienen sentido dentro de una propuesta de innovación o de intervención educativas. Aunque puede argumentarse que los resultados del diagnóstico o de la evaluación no equivalen necesariamente a imperativos de acción, no es menos cierto que una evaluación o un diagnóstico en educación *per se* es algo que carece de sentido. Cuando diagnosticamos o evaluamos perseguimos, aunque sea de una forma indirecta, la mejora de la práctica; en ambas acciones siempre es posible identificar una intencionalidad formativa.

Un tercer y último supuesto que voy a considerar aquí es que la polémica sobre la relación entre diagnóstico y evaluación es el reflejo de los propios cambios que vienen produciéndose en la filosofía, la metodología y la práctica tanto de la evaluación como del diagnóstico. Estos cambios obedecen, a su vez, a modificaciones de naturaleza más profunda que afectan a las estructuras mismas de las sociedades occidentales, al modo en que se introducen e implementan las reformas educativas, y a los propios modelos que tratan de explicar el proceso de enseñanza y aprendizaje.

Tomando como base estos supuestos, entendemos que el diagnóstico está incorporando criterios valorativos en la elección de los procedimientos de mejora de la práctica educativa, y la evaluación está utilizando interpretaciones diversas de la realidad educativa como base para establecer criterios de valor. Diagnóstico y evaluación, pues, han modificado sus planteamientos iniciales y ahora intentan acometer unas tareas cada vez más similares, respondiendo a parecidas finalidades y utilizando prácticamente las mismas estrategias y puntos de referencia conceptuales. La supervivencia de uno de los dos conceptos es una cuestión cultural que escapa a nuestras posibilidades cambiarla. No obstante, en esta polémica la evaluación lleva ventaja.

Claro es que, junto a los nuevos modos de entender el diagnóstico y la evaluación, permanecen otras formas de evaluar y diagnosticar, que al formar parte de una cultura que tiene ya varios decenios posiblemente subsistirán todavía durante mucho tiempo.

Para defender la idea de que evaluación y diagnóstico son conceptualmente similares y que evaluar y diagnosticar son actividades cada vez menos diferentes, vamos a analizar algunos aspectos relativos a ambos conceptos.

1. Contradiendo muchos planteamientos anteriores, la investigación sobre las innovaciones educativas nos revela que el éxito de los cambios descansa al menos tanto en la articulación de un programa técnicamente válido y en el apoyo de instancias académicas y administrativas, como en su respeto a los intereses de determinados grupos sociales y a los valores presentes en determinado contexto cultural.

Quizás el mundo social no sea tan simple, ni la introducción de mejoras educativas se enfrente sólo a las dificultades técnicas inicialmente previstas. La gente no espera las mismas cosas de los programas educativos ni tienen las mismas ideas sobre el modo en que deben desarrollarse. Sus valores y, de hecho, sus intereses difieren. Tampoco todos se ven afectados por las reformas de igual modo y suelen ser distintos sus criterios para juzgar el éxito, mérito o valor de tales reformas.

La evaluación, en unas sociedades cada vez más plurales, se ha visto obligada a cambiar desde posiciones monolíticas a concepciones pluralistas caracterizadas por la utilización de métodos, criterios, perspectivas, audiencias e incluso intereses múltiples. Ha ido abandonando, progresivamente, modelos simplistas que veían la formulación de juicios y la adopción de decisiones como actividades sistemáticas y fácilmente controlables. Al mismo tiempo, la evaluación se acerca cada vez más a concepciones que destacan, como pasos previos para la emisión de juicios, la recogida de una información plural y la identificación de criterios de valor a partir del sondeo de todas las posibles interpretaciones de un hecho.

El diagnóstico, por su parte, se ha visto obligado a modificar su centro de atención desde el sujeto al sujeto y su entorno. Cada vez tiene más sentido la idea de que el diagnóstico no puede establecer explicaciones o realizar pronósticos al margen de las exigencias sociales y escolares que se explicitan en un contexto educativo concreto. Cada vez se invita más a los profesionales del diagnóstico a que conozcan las características del medio escolar, el currículum que se desarrolla, los profesores que lo imparten, los medios con los que se cuenta, etc. También se entiende cada vez más necesaria la implicación de todos los afectados por los resultados del proceso diagnóstico: padres, profesores y alumnos, etc.

Indudablemente, esta «salida al contexto» ha llevado a cuestionar ciertas prácticas educativas (no sólo ciertos aprendizajes) y a entrar en una valoración contextualizada del tipo de intervenciones más apropiadas a las necesidades, valores e intereses de los sujetos. No hay que olvidar que un proceso de evaluación o de diagnóstico puede afectar a más de una persona, a más de grupo social y a más de una institución y cada uno de ellos puede tener diferentes perspectivas, necesidades e intereses relacionados con el modo en que se informa sobre la innovación o la intervención.

2. En las últimas décadas se han producido importantes progresos en el campo de la investigación sobre la enseñanza. Explicar cómo se enseña, por qué se hace de ese modo y cuáles son sus efectos sobre el aprendizaje del alumno, son preguntas que en la actualidad tienen respuestas desde esquemas de trabajo conceptuales más comprensivos.

La investigación ha revelado la importancia de explicar los procesos de enseñanza y aprendizaje teniendo en cuenta variables como los métodos de enseñanza o la conducta docente; los atributos (cogniciones, valores, rasgos personales, etc.) propios de profesores y alumnos así como las interacciones e influencias que ejercen unos sobre otros dentro de la clase; los factores contextuales (ambiente, recursos, titularidad de las instituciones educativas, etc.) y las variaciones contextuales (currículum, asignaturas, materiales de enseñanza...); y los efectos de la enseñanza sobre los alumnos individuales, sobre la clase o sobre las instituciones educativas en su conjunto. Obedeciendo a sus respectivas tradiciones, diagnóstico y evaluación han tratado de integrarse en los nuevos esquemas explicativos adaptando su filosofía, metodología y plan de acción.

El diagnóstico, al igual que la evaluación, ha cedido terreno en sus pretensiones por aportar explicaciones científicas a problemas teóricos en favor de la búsqueda de soluciones a los problemas planteados por la práctica educativa. Diagnóstico y evaluación pueden entenderse como procesos de análisis de la realidad educativa que nos posibilitan conocer cómo debemos articular nuestras intervenciones si queremos mejorarlas.

En términos generales, la finalidad básica a la que responden diagnóstico y evaluación no es otra que la de clarificar problemas, dilemas y tensiones asociados a determinadas prácticas educativas. Contribuyen, de un modo indirecto, al diseño de propuestas de intervención ofreciendo una mejor comprensión de las dificultades, contradicciones y ambigüedades que subyacen en los esfuerzos de mejora de la educación. Diagnóstico y evaluación ofrecen, pues, un filtro que nos sensibiliza hacia ciertas posibilidades de acción y nos predispone en contra de otras.

Buscando una mayor comprensión de la realidad educativa, diagnóstico y evaluación han ampliado su campo de acción hasta cubrir cualquier hecho educativo y sus efectos. El currículum, los programas, las organizaciones educativas, la innovación escolar, etc. son también ahora sus objetos de conocimiento. Lógicamente, la amplitud de su estudio será diferente en función de las metas fijadas en cada intervención y de los propios objetivos que se pretenda cubrir.

3. Diagnóstico y evaluación se desarrollan hoy desde planes de acción más holísticos en los que tienen cabida el análisis de necesidades, el estudio de la implementación de innovaciones e intervenciones educativas, el examen de sus resultados o la elección entre procedimientos de mejora.

Como partes de una propuesta de intervención o de innovación se centran en el *análisis de necesidades*, tanto de los destinatarios como del contexto en que se va a desarrollar la intervención. Pero también diagnóstico y evaluación juegan un papel fundamental en el desarrollo de la intervención, describiendo los hitos (temporalización, materiales, actividades, tareas, implicación de los participantes, reacciones de los participantes, etc.) que nos indican cuál es el *proceso real de ejecución de la misma*. Por último, tras la finalización de la misma, diagnóstico y evaluación permiten valorar el éxito, mérito o valor de la intervención realizada. A partir de los resultados (en los que se pueden detectar nuevas necesidades) es posible que se replanteen tanto las metas como las necesidades inicialmente detectadas, lo que daría lugar a un nuevo ciclo de intervención.

De acuerdo con este planteamiento, entendemos que diagnóstico y evaluación tienen un carácter procesual, que también posee la intervención en la que se integran. No obstante, diagnóstico o evaluación, en cuanto tienen que responder a casuísticas o problemáticas muy diferentes desde muy diversos enfoques, no se acogen a una propuesta única sino que se desarrollan según diversos modelos. Haciendo una abstracción podemos, en cualquier caso, establecer una serie de fases generales: (a)

descripción o de recogida de información; (b) interpretación, que supone dar sentido a la información recogida aportando una explicación razonada desde las propias posiciones del profesional que realiza el diagnóstico o desde los esquemas explicativos que poseen aquellas personas o colectivos afectados¹; (c) valoración, que implica determinar el éxito, mérito o valor de algo comparando la interpretación realizada con unos criterios previamente acordados o fijados a posteriori por el profesional o por los grupos afectados por la innovación.

4. Diagnóstico y evaluación han incorporado nuevos modos de acercarse a la realidad objeto de estudio. Así, si se quería recoger información sobre nuevas variables, respetando una pluralidad de opiniones y valores, había que recurrir junto a los métodos tradicionales a otros ampliamente utilizados en la antropología y la sociología. Tanto el proceso diagnóstico como el de evaluación están abiertos, pues, a la utilización de diferentes diseños: experimentales o cuasi-experimentales; etnografía; investigación-acción; estudios de casos, etc. De igual modo, pueden utilizarse métodos de observación, procedimientos de encuesta (cuestionarios, entrevistas), procedimientos de medida, etc.

En relación con el objeto de estudio en el proceso de diagnóstico o en el de evaluación deberían recogerse dos tipos de informaciones: descriptiva y valorativa. La información descriptiva incluiría, de una parte, las creencias, intenciones y justificaciones, que junto a datos de observación, permitan conocer las condiciones previas a la introducción de la intervención educativa; y, de otra, el desarrollo de la innovación o intervención con los resultados de la misma. La información valorativa estaría compuesta bien por criterios estándares bien por juicios específicos formulados por los afectados por la intervención, o por ambos.

5. El ejercicio profesional del diagnóstico no aparece, en nuestro contexto educativo, claramente delimitado del de la evaluación. La figura del evaluador profesional es prácticamente inexistente. Las actividades de evaluación y las de diagnóstico² son realizadas, según niveles educativos y de intervención (aula, centro, sectores), profesionales tales como profesores, tutores, maestros de educación especial, orientadores, departamentos y unidades de orientación, equipos interdisciplinarios, etc. El diagnóstico del alumno que tiene por objeto proponer modalidades de escolarización —actividad identificada en la nueva regulación del sistema educativo como evaluación psicopedagógica—, es competencia de los equipos multiprofesionales.

1 Los términos y criterios de comparación deben estar establecidos por el profesional o establecerse a lo largo del proceso por los afectados en él.

2 Nos referimos aquí a un diagnóstico comprensivo cuyos resultados se utilizan con otra finalidad que la mera ubicación institucional de los sujetos.

¿POR QUÉ DIFERENCIAR LOS TÉRMINOS *EVALUACIÓN* Y *DIAGNÓSTICO*?

por
Ramón Pérez Juste

1. PLANTEAMIENTO DEL TEMA

Diagnóstico y evaluación son dos conceptos de amplio uso en nuestro marco pedagógico, conceptos que durante mucho tiempo han convivido con sus ámbitos de «competencias» claramente diferenciados, sin superposición alguna.

Esta diferenciación ha venido manifestándose, incluso, en una responsabilización por parte de diferentes Departamentos, primero, y de Áreas de Conocimiento, después.

Sin embargo, de un tiempo a esta parte se están produciendo planteamientos teóricos, con sus correspondientes demandas «prácticas» en ocasiones, que cuestionan explícitamente, o de hecho, la diferenciación entre ambos conceptos¹.

2. POSIBLES RAZONES

Diagnóstico y evaluación son dos conceptos inicialmente nítidos, con un campo específico propio; el *diagnóstico* nace ligado a variables psicobiológicas, tales como las aptitudes, los intereses o el complejo que denominamos *personalidad*, mientras la *evaluación* se liga fundamentalmente a los aprendizajes, con especial incidencia en ámbitos académicos y, dentro de ellos, del denominado *dominio cognoscitivo*.

Por otra parte, también inicialmente, los sujetos a quienes se aplican suelen ser distintos; en el *diagnóstico* se piensa inicialmente en aquellos que tienen problemas o dificultades pues, de hecho, en su acepción genuina el concepto se refiere al «arte descubrir e interpretar los signos de una enfermedad»², mientras en la *evaluación* se trabaja con alumnos «normales».

Del mismo modo, ambos conceptos se mantienen, en general, «limpios», esto es, sin adherencias;

1 Tal vez una de las manifestaciones más palmarias de esta *confusión* pueda ser la que representa, en un campo disciplinar tan próximo a nosotros como la Psicología, la existencia de una asignatura, denominada Psicodiagnóstico, cuyos contenidos se inician con el término *evaluación* en 22 de los 36 temas.

2 Así se define *diagnos* en el Diccionario de uso del español, de María Moliner. En nuestro contexto, «enfermedad» puede entenderse, en sentido amplio, por dificultad, problema, conflicto...

si el *diagnóstico* se concentra en «descubrir e interpretar los signos», la *evaluación* lo hace en «valorar, justipreciar, tasar», algo que, por lo general, se traduce en marcos académicos en términos de *calificación*.

Hasta los momentos son diferentes; el *diagnóstico* suele preceder a las actuaciones de los educadores, sean profesores ordinarios sean especialistas: orienta y guía el tratamiento, que se diseña / elige como el más adecuado para remover las causas de la dificultad o problema identificado. La *evaluación*, por el contrario, se concentra en la etapa posterior al tratamiento o, mejor por lo general, a la actividad de los profesores —enseñanza habitualmentepara apreciar los logros / resultados, en la mayoría de las ocasiones sin otro tipo de trascendencia que promocionar a otra unidad de aprendizaje —o curso académico— o de repetir una u otro.

Los encargados de la actividad son, así mismo, diferentes; el *diagnóstico* es competencia de profesionales que o se hallan fuera de los centros educativos o, si trabajan en ellos, no lo hacen habitualmente como profesores; por contra, la *evaluación* es una de las funciones habituales de todo profesor. Profesor que, ante dificultades o problemas, si le es posible, acude a otros profesionales, entre ellos los encargados, justamente, del diagnóstico.

3. CAMBIOS QUE ORIGINAN LA CONFUSIÓN

A mi juicio se dan dos hechos que trastocan la situación de claridad conceptual, dos hechos, por cierto, reducibles, como veremos, a uno: el estudio *netamente pedagógico* de ambos conceptos.

* El primero se relaciona con los planteamientos teóricos que ligan las intervenciones de los especialistas, en especial de los orientadores, a *TODOS* los educandos en contra de su limitación a los que presentan problemas, dificultades, carencias.

La actuación profesional de los orientadores no se debe restringir ni a los individuos socialmente privilegiados, cuyos padres pueden pagar colegios con servicios especiales —como lo es la orientación en su amplia concepción—, ni a los educandos con problemas; las teorías psicológicas sobre la *individualidad*³, y Su correlato pedagógico de la *educación personalizada*, conducen a considerar tanto que todos los educandos son *especiales* cuanto que, y en coherencia con tal planteamiento, el currículo en general, y el trato pedagógico en particular, deben *adaptarse* a todos y cada uno de ellos.

Pues bien, la *adaptación* —personalización en definitiva— requiere un conocimiento de todas las variables significativas al caso de todos y cada uno de tales educandos.

* El segundo tiene que ver con la apertura del concepto de *evaluación* desde una finalidad y función puramente sancionadora y acreditadora ante la sociedad de los niveles de logro de los alumnos a otra, más pedagógica, auténticamente educativa, como es la denominada *formativa*, que incluye el concepto menos preciso y comprensivo de *evaluación continua*.

De hecho, en mi modesto entender, ambas cuestiones pueden explicarse sencillamente por el hecho de hablar de *diagnóstico* y de *evaluación* no desde el colonialismo de otras ciencias⁴, en especial de la *Psicología* y de la *Medicina*, sino desde la perspectiva, el marco, las funciones y la esencia de la *EDUCACIÓN*, campo específico de la *Pedagogía*. No olvidemos que hablar de *educación* no es sólo hacerlo de los resultados —personas educadas— sino y muy es pecialmente en nuestro ámbito profesional, de *procesos*, en cuyo marco se desdibujan muchas de las diferencias anteriores.

3 Vid. G. W. Allport en *La personalidad. Su configuración y desarrollo*, Barcelona, Herder.

4 Y de otros idiomas y culturas, podríamos añadir.

4. CONSECUENCIAS

Veamos las consecuencias de estos hechos.

Un planteamiento pedagógico netamente personalizado representa la necesidad de una educación caracterizada por un proyecto educativo caracterizado, de una parte, por la **totalidad e integralidad** en sus metas, y, de otra, por su **adecuación** a el educando⁵, concreto e irrepetible, que cada uno es.

Estos tres grandes criterios de calidad se pueden resumir en el concepto de **personalización**, en cuyo marco empieza a desdibujarse la nitidez de la distinción previa entre ambos conceptos.

4.1. Atención orientadora para todos los educandos

Al plantear una demanda de calidad de la educación como una atención personalizada y, por tanto, destinada a **todos** los educandos, porque todos lo necesitan y todos tienen el mismo derecho, el **diagnóstico** deja de centrarse en el «arte de descubrir e interpretar los signos de una enfermedad» para orientarse al **conocimiento** de todos los educados en el conjunto de dimensiones, rasgos, variables... que permiten la personalización tanto del **currículo** como del **trato** personal. Ya no será sólo **correctivo** sino que adquirirá funciones **preventivas**; dejará de ser **reactivo** para pasar a ser **proactivo**; no se centrará en la solución de problemas o en la «**curación**» de «enfermedades» para llegar a incorporar la **potenciación** de las capacidades de **todos y cada** uno de los educandos.

Por otra parte, se procurará que no se encuentren separados los procesos diagnósticos y de tratamiento, aunque éste pueda ser aplicado por profesionales diferentes. Es más, se considera fundamental el **seguimiento** de los «tratamientos» mediante procesos evaluadores, destinados en este caso a apreciar los cambios y a proceder en consecuencia.

4.2. Los cambios hacia una evaluación pedagógica

La evaluación que realizan determinados profesionales se concreta, esencialmente, en el acto de **valorar**; la recogida de información es fundamental, y un momento del proceso, pero éste concluye en el ofrecimiento de la misma, convenientemente valorada⁶ a la persona / institución que la demanda.

Sin embargo, la evaluación plenamente pedagógica, de carácter educativo, debe orientarse a la **mejora** de las personas y no sólo, ni fundamentalmente, al **control** de su aprendizaje y a su **certificación/ sanción social**⁷. Por y para ello no sólo es imprescindible la **valoración** sino que de ésta debe seguirse de la **toma de decisiones**, decisiones que, a su vez, y para tener probabilidades de eficacia y permitir la optimización de los procesos docentes, debe continuar en una fase de **sentimiento**.

Es más, en el marco de esa misma educación de calidad, esto es, **personalizada**, la evaluación no puede limitarse a la constatación de los logros —evaluación sumativa— ni siquiera a proporcionar información procesual de cara a la mejora de los procesos de aprender y, de paso, de los resultados,

5 Vid. al respecto de estos planteamientos el capítulo del autor sobre «Educación de calidad y calidad de la educación», en R. PÉREZ JUSTE, C. MARTÍNEZ MEDIANO y RODRÍGUEZ GÓMEZ (en prensa): *Calidad de los centros educativos*, Madrid, I.T.E.

6 De todos es sabido que, para algunas personas, la **valoración** no es un componente esencial de la evaluación. Vid. al respecto la aportación de D. Nevo, «The Conceptualization of Educational Evaluation» en la obra de HOUSE, *New Directions in Educational Evaluation*, publicada en 1986 por Falmer Press.

7 A nuestro entender, la diferencia se da entre profesionales de la evaluación y profesionales que deben ejercer la evaluación como una función más, junto a otras, al servicio de su finalidad y de los objetivos que les son específicos. En el segundo caso podríamos situar profesionales como el médico o el profesor.

sino que, a fin de acomodar —adecuar, adaptar— los diseños, planes o programas a la diversidad del alumnado, ha de realizar evaluaciones iniciales⁸, de carácter diagnóstico, que permitan anticiparse a las dificultades y problemas, exigir a cada uno según sus posibilidades y hasta descubrir los aspectos en que cada uno de los educandos puede alcanzar metas especialmente importantes de cara a ayudarles a proponerse **programas personales** junto a los objetivos obligatorios y optativos.

Todavía una nueva cuestión; la personalización de los proyectos educativos exige, también, la personalización como referencia a la hora de evaluar. Lo que se ha dado en llamar **rendimiento satisfactorio** tiene repercusiones que hacen más difícil deslindar nuestros dos conceptos; en efecto, la valoración personalizada exige conocer al alumno no ya en sus aprendizajes —campo clásico, y restringido, de la evaluación— sino también en sus capacidades, intereses, motivaciones, condicionamientos familiares... y todas cuantas variables estén implicadas de una u otra forma en los resultados del esfuerzo —estudio— de los alumnos.

Es más, el reconocimiento de que determinadas variables contextuales y técnicas —centros, programas— están implicadas en los resultados del propio aprendizaje académico y, desde luego, en la formación integral de la persona, ha hecho aconsejable **extender** el objeto de la evaluación a centros y programas docentes; pero, en uno y otro caso, se asume que el diagnóstico, en sus facetas *exploratoria e indagatoria*, se encuentra implicado⁹. También, pues, el objeto ampliado pretende ser campo de reflexión de ambos conceptos.

CONCLUSIÓN

Así, al final, las diferencias se diluyen como consecuencia de un proceso de progresivo enriquecimiento y expansión de ambos conceptos; tal vez las únicas que se aprecian con nitidez sean las relativas al **profesional** responsable y, en consecuencia, y ligado a ello, la **profundidad** o **especialidad** en la tarea. El profesional habitual de la evaluación es el profesor, que, por su capacitación, se centra en la apreciación de los resultados del aprendizaje, mientras en el diagnóstico lo es un especialista, al menos en un primer grado, como lo es el orientador, o en un segundo, como los miembros de los equipos psicopedagógicos que, para bien, deberían ser auténticos especialistas en diversos campos (dislexia, adaptación, desarrollo de la inteligencia, sordera, ceguera, lenguaje...).

La atribución de la tarea a diferentes profesionales hace, al final, que el objeto del diagnóstico y el de la evaluación sean diferentes, aunque las tareas que deban realizar sean muy próximas.

También podría pensarse en algo así como la **intención**, si bien no en exclusividad sino en preponderancia: tal intención es más de indagación en el diagnóstico mientras predomina lo valorativo en la evaluación. Pero, en uno y otro caso, la indagación y lo valorativo pueden servir para diseñar intervenciones pedagógicas y/o para apreciar los aspectos que explican resultados satisfactorios, insatisfactorios o mejorables.

8 Dése al término «iniciales» un sentido amplio, capaz de incluir aspectos como el principio de curso o de un período cualquiera e, incluso, el inicio de una nueva unidad de aprendizaje.

9 No olvidemos que se habla de *diagnóstico de necesidades*, en forma más o menos paralela a la evaluación. La Ponencia I de nuestro V Seminario así lo hace.

¿SE EVALÚA Y/O SE DIAGNOSTICA?

por
Ángel Lázaro Martínez
Universidad Complutense Madrid

La pregunta que nos convoca en esta reunión puede parecer tópica y rocambolesca. Es tópica porque, según planteamientos retóricos clásicos, es un lugar común, frecuente y habitual en determinados ambientes y contextos; y por otra, es rocambolesco —calificación aplicada por extensión del famoso personaje novelístico creado por Ponson du Terrail— porque es una cuestión apasionante, audaz e increíble. Y, en consecuencia, la respuesta a semejante cuestión debería estar en los mismos esquemas y planteamientos conceptuales. Por tanto, sin atreverme a profundizar en los laberintos teóricos, intentaré desvelar algunas notas a la, desde mi perspectiva, aparente confusión en la que se mueve la relación diagnóstica-evaluadora en el ámbito pedagógico.

De las diferentes percepciones sobre la similitud o diferenciación del diagnóstico y la evaluación, cabe partir si son términos y procesos identificables, complementarios o dispares. Un primer análisis del constructo, inicial y sencillo, cabe plantearlo desde un mero estudio semiológico y procesual, esto es, constatar cuál es el sentido y finalidad del diagnóstico y de la evaluación en el contexto educativo. Por otra parte, existen otros indicadores que tienden más a confundir que a esclarecer. Por citar un ejemplo, Fernández Ballesteros (1983) señala que la primera referencia histórica sobre evaluación conductual se localiza en el texto de Mischel (1968). Es claro que este autor se refiere a un nuevo enfoque del diagnóstico tradicional respecto al diagnóstico conductual, del que nos ocuparemos más adelante; pero básicamente quisiera destacar que el texto de Mischel, traducido al castellano por primera vez en 1973, con el título de **Personalidad y evaluación** fue publicado originalmente por Wiley y Sons con el de **Personality and assessment**. Estas menciones nos remiten a la comparación de cuatro términos: **evaluación-diagnóstico-assessment** y **evaluation**. Aunque parecen similares, no tienen la misma significación, por lo que conviene apreciar las matizaciones de sus diferencias, para encuadrar la respuesta a la pregunta que nos convoca.

Pero en primer término, vamos a comparar los conceptos que intervienen y componen cada constructo de diagnóstico y evaluación. Tomando como referencia dos trabajos anteriores (Lázaro, 1988 y 1990) se comparan, realizando un breve análisis de contenido, cuarenta definiciones, veinte de cada término, utilizando las concreciones de autores como De la Orden, Ballesteros, Grounland, Lafourcade, Mager, Popham, Provus, Stufelbeam, Tenbrik o Meili, entre otros. Sintetizando aparecen, como confluencia conceptual, los términos de **anamnesis**, **recogida de información**, **exploración**, **medida**, **análisis de datos**, **descripción**, **predicción**, **valoración**, **devolución de información** y **toma de decisiones**. Cada uno de estos tiene referencia distintas, según los enfoques teóricos, pero aparecen como un planteamiento procesual que, básicamente, abarca tres fases claves tanto en el diagnóstico como en la evaluación:

- recoger y sistematizar datos respecto a un hecho
- valorar la información
- tomar decisiones

Esto entraña similitud en el proceso. Entender que el diagnóstico es una mera exploración técnica y una sistematización de datos, excluyendo la planificación de la acción, así como la valoración de los datos y la toma de decisiones, es un enfoque reduccionista. En general, en esta percepción se plantea que la finalidad del diagnóstico es de carácter *descriptivo*, incluso en su relación pronóstica, mientras que la evaluación es un quehacer más amplio pues planifica, valora y toma decisiones de intervención respecto al hecho que se diagnostica. Es un planteamiento que asume e incluye el diagnóstico, y que se permite establecer una *evaluación diagnóstica*.

Pero esta conclusión es, desde nuestra perspectiva, insuficiente y sesgada. Claro que todo se puede aceptar como norma convencional y, en consecuencia, asumir lo que un determinado modelo proponga en un contexto socio-científico predominante. Por tanto cabe entenderlos como diferentes y que existe una cierta relación de dependencia y complementariedad entre el diagnóstico y la evaluación. Sin embargo, sigamos refiriéndonos, en este elemental análisis de los conceptos, al estudio de las definiciones, finalidades y procesos que tanto el diagnóstico como la evaluación realizan.

La primera acepción etimológica del diagnóstico es «conocer a través de» interpretando el prefijo *dia* tanto a nivel temporal como de aplicación de recursos para obtener tal conocimiento. Sencillamente, diagnosticar supone «conocer algo utilizando unos medios a través del tiempo», lo cual permite apreciar la evolución de ese «algo», esto es realizar una «prognosis» o pronóstico. Pero para conseguir esta finalidad es preciso delimitar el problema, el sentido de tal definición, valorar los datos recogidos y, por supuesto, tomar unas decisiones para intervenir. No es concebible, desde la interpretación más clásica de diagnóstico, que solamente se tome una referencia de datos para describir un síndrome sin realizar valoraciones de su intensidad y sus posibles implicaciones en la realidad y, como consecuencia, decidir cuál es la intervención más adecuada. Por tanto, desde esta perspectiva, las pretensiones y procesos del diagnóstico y de la evaluación son identificables, aunque se puedan localizar especializaciones en cada fase del proceso (exploradores, analistas, ..).

Como conclusión cabe entender dos alternativas;

- a.— el diagnóstico incluye la evaluación, como un fase de su proceso.
- b.— La evaluación abarca un amplio proceso, en donde el diagnóstico se reduce a un carácter técnico descriptivo e inicial.

En este caso el enfoque es el que resalta una fase del mismo proceso, aunque no distingue la identidad de la finalidad entre el diagnóstico y la evaluación. Probablemente la diferencia o similitud cabe localizarla por otros análisis. Por ejemplo, desde una perspectiva tradicional, es claramente diferenciable el quehacer diagnóstico del correspondiente al tratamiento consecuente. Pero un diagnóstico evaluación ha superado, en la actualidad, el carácter descriptivo estático de un hecho o de un fenómeno; hay que entender que toda acción diagnóstico-evaluadora en educación, al tomar decisiones, induce a la aplicación de una «intervención mejorativa», lo que requiere un constante «seguimiento», realizando permanentes revisiones diagnósticas para ratificar o rectificar la intervención sugerida. En este enfoque, la evaluación y el diagnóstico adquiere un sentido dinámico, en función de la evolución de un problema en el se está controlando la intervención consecuente de un diagnóstico inicial. Fernández Ballesteros y Carobles (1983) analizan este enriquecimiento diagnóstico-evaluador desde la perspectiva de la conducta Y como referencia a nuestra pregunta inicial, cabría interrogarse si existe una diferencia entre el diagnóstico y la evaluación al incluir la intervención en sus procesos. No obstante, es evidente que, con posibles matices diferenciadores, tanto el diagnóstico como la evaluación, —si fueran diferentes— adquieren, en la actualidad, notas de dinamicidad evolutiva debido al nuevo concepto de conducta. Es más frecuente que se aplique el término «evaluación conductual» que

el de «diagnóstico conductual», tal como señala Silva (1989), dadas la asunción del diagnóstico al «modelo médico o tradicional», distorsión que induce al error entre los mismos especialistas.

Como resumen de lo comentado, el diagnóstico y la evaluación no aparecen, desde mi criterio, como claramente diferentes en su intencionalidad y proceso, incluso en su dinamicidad respecto al seguimiento y control en la intervención consecuenta a la valoración de los datos. Tal vez cabe entender que uno y otro centren su actividad en uno u otro campo a diagnosticar o evaluar.

Pero retomemos a la relación de los términos en inglés y en castellano. Wolf (1990) señala que «assessment» se refiere a personas mientras que «evaluation» debe aplicarse a entidades abstractas, como programas, curricula y variables organizativas. Estas apreciaciones nos remiten a la relación entre los términos ingleses (assessment, evaluation, appraisal, diagnosis) y castellanos. A veces, se tiende a confundir la evaluación inglesa con evaluación lo cual puede provocar confusiones, debidos a las matices de los constructos en un idioma u otro. Wegner (1983) señala

evaluación es un concepto tomado del inglés y que ha alcanzado amplia aceptación en la literatura psicopedagógica de otras lenguas (458)

De una forma habitual, en la Enciclopedia Internacional de Evaluación educativa (1990), se diferencia «assessment» y «evaluation», (Choppin), entendiendo que lo primero se refiere a los procedimientos y procesos utilizados para recoger datos relativos a la conducta humana; y, reiterando lo indicado anteriormente, «evaluation» se aplica a entidades abstractas. En este caso la diferencia se refiere al objeto de estudio, entendiendo que el proceso y la intencionalidad es la misma. De esta forma, el diagnóstico es una fase, tanto de la evaluación como del assessment. Y así cabe entender que el diagnóstico se aplique tanto a personas como a instituciones.

Como conclusión de lo expuesto, las opciones que actualmente provocan polémicas entre el diagnóstico y la evaluación son dependientes del enfoque técnico y del modelo científico que se asuma. Fundamentalmente, desde mi perspectiva, con muchas matizaciones conceptuales, entendemos que el hecho diagnóstico implica la delimitación dinámica de situaciones educativas, para valorarlas y tomar decisiones de intervención que afecten a alumnos, profesores, programas, procesos, productos, organizaciones, instituciones o sistemas. Y, por tanto, al diagnosticar se describe y se evalúa; y, por supuesto, el se evalúa porque se está diagnosticando.

REFERENCIAS

- CHOPPIN, B. H. (1990): «Evaluation, assessment and measurement», en Encyclopedia International of evaluation educational. Pergamon Press. Oxford.
- FERNÁNDEZ BALLESTEROS, R. (1983): Evaluación Conductual. Pirámide. Madrid.
- FERNÁNDEZ BALLESTEROS, R y CARROBLES, J. A. (1983): «Comparaciones entre evaluación tradicional y evaluación conductual» en Evaluación conductual. Pirámide. Madrid.
- HAETEL, G. D. y WALBERG, H. (1990): The international encyclopedia of educational evaluation. Pergamon Press. Oxford.
- LÁZARO, A. (1987): Diagnóstico Pedagógico. Proyecto docente. Madrid. Universidad Complutense.
- LÁZARO, A. (1990): «Problemas y polémicas en torno al diagnóstico pedagógico». Bordón, 40, 1, 7-15.
- MISCHEL, W. (1980): Personalidad y evaluación. Trillas. México.
- SILVA, F. (1989): Evaluación conductual y criterios psicométricos. Pirámide. Madrid.
- WEGENER, W. (1983): «Evaluación», en Rombarh, H. (ed.): Diccionario Ciencias de Educación. Rioduero. Madrid.
- WOLF, R. M. (1990): «The nature of educational evaluation», en The international encyclopedia of educational evaluation. Pergamon Press. Oxford.

COMUNICACIONES AL SIMPOSIUM I

DISEÑO Y PLANIFICACIÓN DE UN PROGRAMA DE INTERVENCIÓN EN ORIENTACIÓN UNIVERSITARIA

Francisca Castellano Moreno y José Antonio Delgado Sánchez
Universidad de Granada

INTRODUCCIÓN

Conceptualizar la orientación como un proceso integrado a lo largo de todo el ciclo vital del alumno, ha llevado a considerar la intervención mediante programas como una estrategia adecuada que da sentido al carácter educativo de la orientación. En dicha intervención, «el análisis o evaluación de necesidades» se constituye en una fase esencial de la planificación del programa para que éste responda a las expectativas de los usuarios (Rodríguez Espinar, 1985; Sanz Oro, 1990; Tejedor, 1990; Collison, 1982).

En la Universidad de Granada, a través del Vicerrectorado de Estudiantes, se vienen desarrollando curso tras curso «jornadas de acogida al estudiante». A su vez, el ICE de esta universidad viene organizando cursos sobre «técnicas de estudio» para los alumnos que ingresan en la universidad. Asimismo, algunas facultades, y a través de sus respectivos decanatos, recogen estas iniciativas y las implementa con su alumnado.

Creada la Facultad de Ciencias de la Educación, que refunde los estudios de la antigua Escuela Universitaria de Formación del Profesorado de EGB y los de la Sección de Pedagogía de la Facultad de Filosofía y Letras, pretendemos aglutinar todas estas iniciativas a través del diseño de un programa de orientación al estudiante como respuesta a sus demandas.

1. ANÁLISIS O EVALUACIÓN DE NECESIDADES

Teniendo en cuenta que este programa va dirigido a estudiantes de primero de la recién creada Facultad de Ciencias de la Educación de la Universidad de Granada, y considerando, por un lado, los cambios que de cara al alumnado están produciéndose en el sistema universitario en general, tales como la implantación de nuevos títulos, los cambios en los planes de estudio, el nuevo sistema de créditos, y la ampliación de la oferta de asignaturas optativas; por otro, los servicios que ofrece el Vicerrectorado de Estudiantes y que el alumno desconoce al ingresar en la universidad; y por último, en el plano de lo concreto, el análisis o evaluación de las necesidades de los estudiantes en orientación llevado a cabo en esta universidad, extraemos aquéllas que en función del tiempo y de los recursos, van a ser atendidas en el programa. Estas son:

- a) Necesidades de carácter informativo:

- El proceso de matriculación y la búsqueda de información de carácter académico.
 - La financiación de sus estudios y su participación en la comunidad universitaria.
 - Los nuevos títulos y planes de estudio, la opcionalidad, y el sistema de créditos.
- b) Necesidades de carácter formativo:
- Adquirir técnicas que le permitan desarrollar con efectividad su función como estudiante

2. ESTRUCTURA DEL PROGRAMA

El programa se estructura en torno a los siguientes elementos:

2.1. Metas y objetivos

En el programa, las metas, referentes últimos del mismo, expresan el rol que éste está llamado a desempeñar. A través del mismo pretendemos:

«Favorecer la integración del alumno universitario en el contexto de la comunidad educativa en el que se va a desenvolver».

Esta meta se traduce en intenciones educativas más concretas, precisando los comportamientos de los alumnos, función ésta que se canaliza mediante el establecimiento de los objetivos. Formulamos los siguientes:

- a. Conocer los principales servicios tanto académicos como administrativos de la universidad más directamente relacionados con sus estudios.
- b. Conectar al alumno con el vicerrectorado de estudiantes para conocer su estructura y los servicios que presta.
- c. Familiarizarlo con su nueva facultad, mediante el conocimiento de su estructura, organización y funcionamiento.
- d. Presentar a los alumnos los nuevos planes de estudio como medio para favorecer la construcción de su carrera profesional.
- e. Equipar al alumno de técnicas de estudio específicas que le faciliten un adecuado progreso en las materias y especialidades elegidas.

2.2. Contenidos

El programa se estructura en torno a los siguientes contenidos:

- La Universidad de Granada: servicios académicos y administrativos.
- El Vicerrectorado de Estudiantes: estructura y servicios.
- La Facultad de Ciencias de la Educación: estructura, organización y funcionamiento.
- Los planes de estudio: estructura, contenido, especialidades y opcionalidad.
- Técnicas de trabajo intelectual

2.3. Estrategias y actividades

Para llevar a cabo el programa estableceremos una estrategia adecuada mediante el desarrollo de las siguientes acciones:

Conectar con los diferentes profesionales que van a intervenir en el programa, y centros de apoyo (Vicerrectorado de estudiantes, ICE, decanato, departamentos).

Adecuar los tiempos y espacios a las diferentes actividades.

Supervisar la disponibilidad de los recursos existentes y elaborar otros necesarios.

Búsqueda de apoyos y recursos financieros.

El programa va destinado a los alumnos de primer curso, estructurándolo mediante actividades de sesiones informativas, mesas redondas, cursos, conferencias. En el mismo se contempla la impartición de contenidos teóricos con la simultaneidad de sesiones prácticas. Para su desarrollo conformaremos los grupos de alumnos en función del tipo de actividad a realizar.

2.4. Temporalización

El programa se diseña para ser llevado a cabo durante los meses de octubre y noviembre del curso escolar 1993-94, en sesiones de mañana y tarde.

2.5. Recursos

Para llevar a cabo el programa contamos con recursos tanto humanos como materiales. Así:

Humanos: Decano, profesores, personal a cargo de los servicios (bibliotecarios, administrativos, de informática, del vicerrectorado), alumnos de prácticas y antiguos alumnos.

Materiales: Folletos informativos, paneles, material de técnicas de estudio, boletines informativos.

2.6. Evaluación

En el programa contemplamos una evaluación interna orientada en dos direcciones:

a) La autoevaluación por los profesores responsables de llevarlo a cabo.

b) La evaluación de los objetivos contemplados en el programa mediante la satisfacción de los usuarios del mismo (los alumnos).

La finalidad de la evaluación será de carácter formativo, buscando la mejora del programa para su implementación en el próximo curso.

BIBLIOGRAFÍA

COLLISON, B. B. (1982): «Needs assessment for guidance program planning: A procedure». *School Counselor*, 30, 115-121.

RODRÍGUEZ ESPINAR, S. (1985): *Proyecto docente e investigador. Memoria para la cátedra de orientación educativa* (inédita). Universidad de Barcelona.

SANZ ORO, R. (1990): *Evaluación de programas en orientación educativa*. Madrid, Ediciones Pirámide.

TEJEDOR, F. J. (1990): «Perspectiva metodológica del diagnóstico y evaluación de necesidades en el ámbito educativo». *RIE*, Vol. 8 (16), 15-37.

PAUTAS PARA LA EVALUACIÓN DE UN PROGRAMA DE ORIENTACIÓN

J. Hernández Fernández y P. Martínez Clares
Universidad de Murcia

INTRODUCCIÓN

El estudio de la práctica orientadora nos revela la insuficiencia de una evaluación sistemática y propia en el campo de la Orientación Educativa. Este hecho se ha atribuido, entre otras razones, al carácter relativamente novedoso de esta praxis, a la resistencia de los orientadores para evaluar su propia actividad, provocada en un principio por el carácter sancionador que adoptaron las propuestas de evaluación, por lo que argumentan que ésta va dirigida generalmente al cambio de conducta y el que los resultados de la misma son en muchas ocasiones sólo valorables a medio y largo plazo, lo que supone una dificultad añadida, pues escapan de todo tipo de análisis o evaluación posible.

Sin embargo, el convencimiento de que es necesario dotar a la Orientación Educativa de una metodología precisa y propia de evaluación que demuestre la eficacia de la intervención y la corrección del proceso de actuación (Humes, 1972; Krumboltz, 1974; Pine, 1975; Knapper, 1978) ha desembocado en la adaptación de modelos didácticos y/o sociales de evaluación al ámbito de la Orientación Educativa (Alkin, 1969; Campbell, 1971; Lewis, 1983; Stufflebeam, 1989) y a la formulación de modelos propios, como el proactivo de Atkinson (1985), que trata de superar las limitaciones de los modelos reactivos.

A pesar de ello las dificultades que los orientadores encuentran cuando pretenden evaluar su propia actuación y logros son numerosas, entre otros motivos, por la ausencia de propuestas y guías específicas que les ayuden a salvar la distancia entre los modelos teóricos y la realidad de su actuación.

Es en esta línea de aproximación y/o facilitación de la práctica evaluadora en donde se puede ubicar la intención de la presente aportación. Para lo que partimos, entre otros supuestos básicos, del entendimiento de que el proceso de evaluación y el de programación/planificación se interrelacionan y se deben producir paralelamente, es decir, la evaluación puede realizarse en cualquier etapa o fase de la programación, produciéndose un proceso de retroalimentación permanente.

Por ello, uno de los objetivos de este trabajo es el de delimitar las fases del proceso de evaluación de programas en el contexto de la Orientación Educativa, a partir del paralelismo, ya señalado, entre éstas y las fases de planificación/programación del programa de orientación propiamente dicho.

De otra parte damos por sentado que el programa se basa en la identificación y determinación de necesidades de la comunidad educativa, a partir de una evaluación del contexto socio-ambiental y educativo del centro escolar donde se va a desarrollar. A la vez que consideramos como finalidad

general de la evaluación de programas de Orientación la toma de decisiones para la mejora del proceso de intervención y del logro de resultados.

FASES EN EL PROCESO DE EVALUACIÓN DE PROGRAMAS DE ORIENTACIÓN EDUCATIVA

Antes de iniciar el diseño de la evaluación de cualquier programa de Orientación (y por tanto en el momento de iniciar la propia planificación y diseño del mismo) hay que fijar los **supuestos teóricos** previos en los que se va a basar dicha evaluación. En ellos se debe especificar la finalidad, propósitos, tipo de información que se va a dar y receptor/es de la misma. Estos aspectos, junto con las características del programa y las del contexto en que se va a llevar a cabo el mismo, nos servirán para determinar el modelo y tipo de evaluación en el diseño de ésta, una vez que se valide el programa y podamos pasar a precisar la metodología de evaluación.

La primera fase, a la que denominamos **conocimiento del programa de Orientación**, permite al evaluador interno examinar la adecuación del diseño del programa a los requisitos mínimos de programación y al evaluador externo familiarizarse y conocer el programa que va a evaluar. Entre los aspectos a analizar se encuentran: el *tipo de programa* —si es de orientación escolar, personal, vocacional, de desarrollo personal y social, etc.—, los *destinatarios* del mismo —si se dirige a alumnos, padres, profesores, etc.—, las *metas y objetivos* del programa. En esta fase deberá conocerse también las *acciones o actividades*, la *metodología* y la *temporalización* del programa.¹ Todo ello nos permitirá conocer «cómo», para «cuándo» y «con qué» medios y recursos se ha planificado la intervención orientadora.

En resumen, esta primera fase del proceso evaluador comprenderá el conocimiento y familiarización de todos los aspectos y factores que configuran el diseño de planificación del programa de Orientación. De ahí que sea necesario que el programa a evaluar esté escrito, con especificación de los aspectos citados.

La segunda fase la llamamos **validación de la evaluación**, también denominada por otros autores como evaluabilidad de la evaluación (Wholey, 1977; Horst, Scanlon y Wholey, 1977). En ella se pretende analizar si es posible la evaluación de un programa, es decir, si el programa responde o no a las necesidades para las que ha sido diseñado, si necesita modificación en alguna o en todas sus partes o debería de ser retirado (Hernández y Rubio, 1992).

Smith (1989) define la evaluabilidad de la evaluación como «*una herramienta prescriptiva y de diagnóstico para mejorar los programas y tomar de decisiones más útiles*». La validación o evaluabilidad de la evaluación se podría utilizar, en palabras del propio Smith, como un «instrumento de la planificación» para consensuar objetivos, identificar actividades, y localizar recursos y medios necesarios para la implantación del programa. Todo esto supone una primera pre-evaluación o toma de decisión acerca del programa de Orientación, pues en esta fase se produce la primera retroalimentación sobre la validez del programa propuesto y sobre la necesidad y conveniencia de realizar esta evaluación. La **validación de la evaluación** nos proporcionará un primer juicio acerca de la suficiencia y adecuación del programa para satisfacer las necesidades o carencias detectadas en la fase previa de Análisis o Evaluación de Necesidades. Con ello se obtiene de una parte, la concreción de las coordina-

1 Empleamos el término **meta** en el sentido en que lo hace Kaufman (1977) de determinación de «*hacia dónde vamos, cómo sabremos cuándo hemos llegado y cuales son las principales etapas para ir del punto en que estamos a donde debemos estar*». Se trata, según el modelo sistémico de Kaufman, de conocer en este momento el *análisis de misiones*, para pasar al *análisis de funciones* (etapas), que determina «qué» debe hacerse para conseguir unos resultados u objetivos factibles para satisfacer las necesidades detectadas.

nadas espacio/temporales y de los costos del programa, y de otra suministrar información del programa a los responsables del mismo; en definitiva esta fase intenta garantizar el éxito del programa de intervención orientadora en la comunidad educativa.

Para realizar esta evaluación del diseño del programa hay que fijar una serie de dimensiones o interrogantes a evaluar. Proponemos las indicadas por Fernández Ballesteros (1988) y Fernández Ballesteros y Hernández (1989): evaluación de necesidades, especificación de objetivos, calidad en los criterios de selección, definición del programa, implantación, calidad del diseño, calidad de las operaciones, procedimientos de recogida de información, posibilidad de recogida de información sobre el contexto, aceptabilidad, obstáculos, implicación del evaluador, finalidad de los resultados evaluativos y coste de la evaluación.

Pero estas dimensiones no quedan determinadas ni concretadas si no se establecen una serie de criterios de validación de evaluación; sugerimos que al menos deberían seguirse los siguientes para la validación de un programa de Orientación, que no podrá ser evaluado si:

1. Sus objetivos no están formulados de forma operativa/conductual (si no son medibles y observables).
2. No tiene delimitadas las acciones o actividades a realizar en unas coordenadas espacio/temporales.
3. No puede predecir los niveles de logro aceptados por la comunidad educativa.
4. No se conocen los recursos materiales y humanos disponibles para implantar el programa.

Lo ideal sería que de cada dimensión antes referida se fijara al menos un criterio o indicador acerca de la validez de la evaluación. De acuerdo con Hernández y Rubio (1992), el peso de cada criterio depende de dos factores: las concepciones y técnicas del evaluador, y las peculiaridades específicas del contexto y personas implicadas en la implementación del programa.

La fase concluye cuando se definen estos criterios y se comprueba que el programa a evaluar cumple con los objetivos de evaluación mínimos que se pueden resumir en que el contexto es favorable para realizar la evaluación, el programa de orientación en sí mismo es evaluable y el evaluador (externo, interno o mixto) es capaz de realizar dicha función.

Es en este momento cuando comienza el diseño de evaluación propiamente dicho y cuando se ha de decidir el tipo de evaluación o qué aspectos del programa se va a evaluar. Aunque un programa de Orientación, como programa socio-educativo, puede ser sometido a distintos tipos de evaluación según sea el aspecto o pregunta que se quiera responder (Rodríguez Espinar, 1986:373), dadas las características de la mayoría de las intervenciones orientadoras desarrolladas en los centros docentes de nuestro contexto y siguiendo una secuencialización jerárquica de los distintos tipos de evaluación (Alvira, 1991), proponemos dos tipos básicos de evolución para los programas de intervención orientadora: **evaluación del desarrollo del programa**, que englobaría a su vez la *evaluación de la implementación* o puesta en marcha y la *evaluación del proceso*, y la **evaluación del producto** o de eficacia, a las que consideramos fases necesarias para una correcta evaluación de un programa de intervención orientadora.

De ahí que la fase siguiente sea la **evaluación del desarrollo del programa**, que tiene como finalidad la recogida de datos y de información de las etapas de planificación del programa a evaluar y de las primeras fases del proceso de evaluación, para posteriormente realizar un análisis de los datos y poder tomar decisiones. El análisis girará en torno a la marcha del programa tanto en su adecuación a la programación previa como a los aspectos dinámicos y de relación del mismo. Tiene por tanto un carácter sumativo y implica la realización de evaluaciones intermedias que permitan la retroalimentación del programa, Pero en cada una de ellas se puede distinguir como fase sucesivas la evaluación de la implementación y la evaluación del proceso.

La **evaluación de la implementación** consiste en evaluar «qué» está funcionando del programa

una vez que se ha puesto en marcha, es decir, la instrumentalización del programa de intervención, su puesta en práctica siguiendo las etapas y esquemas teóricos previamente concretados. El fin último de esta fase de evaluación, es contrastar si hay o no discrepancias entre el diseño y la realidad; y en caso afirmativo, realizar la adaptación pertinente, redefinir el programa para lograr su óptima y adecuada puesta en marcha.

En ella se debe recoger información sobre el cumplimiento del programa en su cobertura, realización de actividades, ejecución de la temporalización implicación y cumplimiento de funciones por los agentes, utilización de recursos... Para ello se debe tener muy claro cuáles son los aspectos nucleares del programa y las unidades de análisis a utilizar (muestreo si es el caso) de modo que se pueda realizar correctamente la comparación entre lo planificado y lo realizado porque en resumen, en esta fase se quiere ver cómo se desenvuelve en la praxis y realidad diarias, el programa de intervención orientadora diseñado previamente, comparando el modelo ideal con el modelo real que se desarrolla.

Lo que realmente interesa en esta fase es *evaluar la identificación y adecuación de las actividades* diseñadas al contexto real, responder a la pregunta de si existe o no actividades uniformes implementadas de manera sistemática (Alvira, 1991); en definitiva queremos evaluar si las actividades que se están implementando son las adecuadas en las coordenadas espacio/temporales establecidas y se pueden desarrollar con los instrumentos, medios y recursos prefijados.

Si la evaluación de la implementación se refiere a «qué» se hace en un programa, la **evaluación de proceso** explica el «como» (Municio,1992:389); de ahí que pensemos que la evaluación de proceso viene a cubrir en gran parte el binomio programa-factor humano, tanto en la faceta del modo de realización de actividades o tareas como en la de las realizaciones que se establecen entre los miembros implicados en el programa (agentes y sujetos). En este sentido la evaluación de proceso nos debe de indicar la adecuación de la actuación (comportamientos) de la actuación de los agentes del programa, tanto a los objetivos del mismo como a los estilos de actuación y metodologías previstos, así como las respuestas de los sujetos del programa a dichos comportamiento. En cierto modo implicaría, entre otros aspectos una autoevaluación del orientador o agente.

De otro lado, la evaluación de proceso debe analizar la adecuación del formato técnico del programa (estrategias, recursos, actividades, tareas...) a los objetivos del programa, en tanto que su realización ayude al logro de los mismos, así como el papel real que dicho formato técnico juega en la realización del programa, en la línea de la evaluación de pertinencia o propiedad y de idoneidad o suficiencia que señalan Aguilar y Ander-Egg (1992:48). En otras palabras, se trata de analizar cómo se han instrumentalizado y utilizado los recursos técnicas y procedimientos del programa.

Para ello es necesario además de la formulación de criterios e indicadores que nos señalen el modo y adecuación del proceso (agentes, relaciones y formato técnico), el contar con un apropiado sistema de control e información sobre la marcha del programa que permita revisiones periódicas del mismo.

Por fin, la última fase de nuestro modelo, la **evaluación de la eficacia y efectividad** del programa de Orientación, trata de realizar una evaluación de resultados y efectos del programa, es decir, consiste como expresa Aguilar y Ander-Egg (1992) en comprobar en qué medida se están alcanzando los resultados previstos y también los efectos y resultados no previstos. En ella se trata de recoger suficiente información para saber si el programa está o no consiguiendo los resultados esperados, a través de dos preguntas principalmente:

1. ¿Se están consiguiendo los resultados esperados?
2. ¿Estos resultados guardan relación y concordancia con el programa diseñado?, es decir, ¿los resultados obtenidos son efecto del programa desarrollado o de otras variables o aspectos no contemplados?

Para contestar «eficazmente» a ellas, los objetivos y resultados esperados deben estar perfectamente identificados y tener una atribución unívoca, de modo que se pueda establecer fácilmente el nivel de

logro alcanzado en la consecución de los objetivos, si el mismo es debido al diseño, implementación y proceso del programa y si se han obtenido otros efectos además de los previstos.

En definitiva, lo que estamos proponiendo en esta última fase es no sólo el análisis del logro de objetivos, sino una evaluación global del programa a partir de establecer la relación entre la consecución de resultados y las posibles causas de los mismos, ya que en ocasiones la mera evaluación de la eficacia puede resultar un dato pobre para la toma de decisiones sobre el programa.

Por ello puede ser necesario incluir otros tipos de evaluación que completen la de resultados. Así se puede realizar una evaluación del *impacto* de los resultados en el contexto en el que se desarrolla el programa y sobre los otros subsistemas a los que no iba dirigido el mismo; o bien una evaluación de *efectos colaterales* o de efectos no previstos, e incluso la evaluación de la *eficiencia* basada en la relación coste-efectividad, coste-utilidad del programa, que son las más provechosas en Orientación, para lo que podemos establecer una serie de indicadores cuantitativos que nos permitirán evaluar los resultados del programa, como por ejemplo indicadores de demanda, de participación, de fondos,...

La posible realización de éstos u otros tipos de evaluación se deben hacer de modo subsidiario a la finalidad de esta fase de evaluación, que no es otra que comprobar el logro de los objetivos y los efectos del programa y las posibles causas de los mismos

Para la realización de la evaluación una dificultad añadida a las que señalamos al principio, es la escasez de instrumentos y las limitadas garantías de fiabilidad y validez de los existentes (García Ramos, 1992), unida a la necesidad de adaptar los mismos a los propósitos de la evaluación, al contexto y a las características del programa. Por lo que proponemos para la realización de este modelo la construcción de recursos metodológicos propios, adaptados al contexto donde surge y posteriormente se desarrolla el programa. La elaboración de estos permitirá ponderar en aquellos aspectos más significativos para nuestro caso particular o concreto y poder recoger la información necesaria para nuestro fin.

Así, disponemos de varios métodos cualitativos como pueden ser: la observación directa, la observación participante, las entrevistas (estructuradas y semiestructuradas), el análisis de contenido (de contexto), la documentación existente sobre el programa, las reuniones de grupo, el estudio de casos,...etc. A través de estos métodos, podemos elaborar instrumentos de evaluación como son las rejillas de observación y/o listas de control y estimación, cuestionarios, donde incluyamos los criterios e indicadores que hemos fijado para nuestra evaluación; como ejemplos de guías para su elaboración podemos citar el Listado de Cuestiones Relevantes en Evaluación de Programas (LCREP) de Fernández Ballesteros y Hernández (1989) y la rejilla de evaluación de programas de Rodríguez Moreno (1988).

CONCLUSIÓN-RESUMEN

El modelo de evaluación propuesto, destinado a la toma de decisiones de programas de intervención orientadora, pretende ser un esquema secuenciado y lógico, que permita a los orientadores, siguiendo la ya clásica afirmación de Herr (1976) demostrar que se ha hecho lo que se quería hacer, a lo que nosotras añadimos, de la mejor manera posible. Para ello el modelo propone las diferentes fases, que responden a las siguientes preguntas básicas:

1. ¿Por qué y para quien evaluamos?
2. ¿Qué programa vamos a evaluar?
3. ¿Se puede evaluar ese programa?
4. ¿Se ha llevado a cabo el programa tal y como se diseñó?
5. ¿Cómo ha sido esa puesta en marcha y por qué funciona el programa?
6. ¿Se ha logrado lo que nos proponíamos y por qué?

MODELO DE EVALUACIÓN DE PROGRAMAS DE ORIENTACIÓN EDUCATIVA

FASES	ASPECTOS A EVALUAR	CUESTIÓN
1. SUPUESTOS TEÓRICOS	<ul style="list-style-type: none"> • Finalidad • Propósitos • Tipo de información • Receptor/es • Caracteres del profesorado y del contexto 	¿Por qué y para quién evaluamos?
2. CONOCIMIENTO DEL PROGRAMA	<ul style="list-style-type: none"> • Tipo de programa • Destinatarios • Conocer metas y objetivos • Conocer las actividades • Temporalización • Recursos y medios... 	¿Qué programa vamos a evaluar?
2. VALIDACIÓN DE LA EVALUACIÓN	<ul style="list-style-type: none"> • Dimensiones o interrogantes a evaluar de Fernández Ballesteros y Hernández (1989): <ul style="list-style-type: none"> * Evaluación de necesidades, especificación de objetivos, calidad de criterios de selección, definición del programa, implantación, calidad del diseño, calidad de las operaciones, procedimientos de recogida de información, aceptabilidad, obstáculos, implicación del evaluador, finalidades de los resultados y coste de la evaluación. * Criterios de evaluación. 	¿Se puede evaluar ese programa?
4. EVALUACIÓN DEL DESARROLLO	<p><i>A. Evaluación de la implementación</i></p> <ul style="list-style-type: none"> • Cobertura del profesorado • Realización de actividades (identificación y adecuación) • Ejecución de la temporalización • Funciones de los agentes • Utilización de Recursos <p><i>B. Evaluación del proceso</i></p> <ul style="list-style-type: none"> • Binomio programa/factor humano • Autoevaluación del orientador o agente • Análisis técnico del programa 	<p>¿Se ha llevado el programa a la práctica tal y como se diseñó?</p> <p>¿Cómo ha sido su puesta en marcha y por qué funciona el programa?</p>
5. EVALUACIÓN DE LA EFICACIA Y EFECTIVIDAD DEL PROGRAMA	<ul style="list-style-type: none"> • Evaluación de resultados y efectos. • Evaluación global del profesor a través de: <ul style="list-style-type: none"> * Evaluación de impacto. * Evaluación de efectos colaterales. * Evaluación de la eficiencia: <ul style="list-style-type: none"> — Análisis coste/efectividad — Análisis costo/utilidad * Indicadores cuantitativos: <ul style="list-style-type: none"> — índice de demanda — índice de participación 	¿Se ha logrado lo que nos proponíamos? ¿Por qué?

REFERENCIAS BIBLIOGRÁFICAS

- AGUILAR, M^a. J. y ANDER-EGG, E. (1992): Evaluación de servicios y programas sociales. Madrid. Siglo XXI.
- ALKIN, M. C. (1969): Evaluation theory development. *Evaluation Corriment*, 2, 2-7.
- ALVIRA, F. (1991): Metodología de evaluación de programas. Madrid. Cuadernos Metodológicos. CIS.
- CAMPBELL, R. E. y otros (1971): The systems approach: An emergin behavioral model for career guidance. Ohio, The Center for Vocational and Thechnical Education.
- FERNÁNDEZ BALLESTEROS, R. (1988): Listado de Cuestiones Relevantes en Evaluación de Programas (LCREP). Madrid. Inédito.
- FERNÁNDEZ BALLESTEROS, R. y HERNÁNDEZ, J. M. (1989): Listado de Cuestiones Relevantes en Evaluación de Programas (LCREP). Madrid. Inédito.
- GARCÍA RAMOS, J. M. (1992): Recursos metodológicos de la evaluación de programas. *Bordon* 43, 461-475.
- HAUFFMAN, R. E. (1977): Planificación de sistemas educativos. México. Trillas
- HERNÁNDEZ, J. M^a y RUBIO, V. J. (1992): Análisis de la evaluabilidad: paso previo de la evaluación de programas. *Bordon* 43, 397-405.
- HERR, E. L. (1976): Counseling: Accountability, reality, credibility. *J. of Counseling Services*, 1, 14-23.
- HORST, P., NAY, J. N., SCANLON, J. W. y WHOLEY, J. S. (1974): Program management and the federal evaluator. *Public Administration Review*, 34, 300-308.
- HUMES, C. W. (1972): Accountability: a boon to Guidance. *Personnel & Guidance J.* 51, 21-26
- KNAPPER, E. Q. (1978): Counsselor Accountability. *Personnel & Guidance J.* 57, 27-30.
- KRUMBOLTZ, J. D. (1974): An accountability Model for Counselors. *Personnel & Guidance J.* 52, 639-646.
- LEWIS, J. D. (1983): Guidance program evaluation-how to do it. *School Counselor*, 31, 111-119.
- MUNICIO, P. (1992): La evaluación segmentada de los programas. *Bordon* 43, 4. 375-395.
- PINE, G. (1975): Evaluating School Counseling Programs: Retrospect and Prospect. *Measurement & Evaluation in Guidance*, 8, 136-144.
- RODRÍGUEZ ESPINAR, S. (1986): *Proyecto Docente e Investigador*. Presentado al concurso de acceso a la plaza de Catedrático de la Universidad de Barcelona en Área de Métodos de Investigación y Diagnóstico en Educación (Orientación Educativa. Barcelona. Inédito.
- RODRÍGUEZ MORENO, M. L. (1988): *Orientación Educativa*. Barcelona CEAC.
- SMITH, M. F. (1989): Evaluability Assessment. A practical approach. Boston. Kluwer.
- STUFFLEBEAM, D. L. y SHINKFIELD, A. J. (1987): Evaluación Sistémica. Barcelona. Nec/Paidos.

LA SATISFACCIÓN/INSATISFACCIÓN EN EL TRABAJO DE LOS ORIENTADORES ESCOLARES: IMPLICACIONES PROFESIONALES

Luis Sobrado Fernández y Ana M^a Porto Castro
Universidad de Santiago

INTRODUCCIÓN

La satisfacción/insatisfacción en el trabajo es un conjunto de sentimientos favorables o desfavorables a través de los cuales los profesionales perciben su actividad laboral. Es parte integrante de la satisfacción vital de la persona. La naturaleza del ambiente externo del sujeto influye en los sentimientos que tenga del empleo, y de un modo semejante al ser la profesión una parte importante de la vida de la persona, la satisfacción en la actividad profesional incide en su satisfacción vital.

Muy relacionada con la satisfacción profesional es el clima de trabajo en las organizaciones sociolaborales que tiene una repercusión directa sobre el rendimiento y la satisfacción de los sujetos en el empleo. El ambiente de trabajo se halla ligado a la situación organizativa y a las percepciones en el mismo; la satisfacción implica una estimación de las condiciones del trabajo o de la organización sociolaboral. Es pues este el aspecto afectivo de la percepción personal, y por ello los miembros de una organización socioocupacional que tienen una buena percepción de su clima laboral tienden a estar más satisfechos en su trabajo (Brunet, 1987).

Según Davis y Newstroum (1990) la satisfacción profesional está relacionada con tres variables principales referidas al trabajador y al ambiente laboral: la edad, el nivel de formación y el tamaño de la organización.

Respecto a la edad cuando los profesionales van aumentando en ella se sienten más satisfechos con su empleo, debido a sus menores expectativas y a un mejor ajuste a su situación laboral, dada su experiencia, excepto cuando se hallan próximos a su jubilación.

Por otra parte, en general los trabajadores que poseen mayores niveles de formación y especialización tienden a estar más satisfechos en sus empleos, porque suelen estar mejor pagados, tienen mejores condiciones laborales y hay una utilización más completa de sus habilidades.

En tercer lugar, el tamaño de la organización sociolaboral, está con frecuencia inversamente relacionado con la satisfacción en el trabajo, de modo que conforme aquélla va creciendo, la satisfacción en el empleo declina, debido a la carencia de contactos personales, menor motivación, comunicación, coordinación y participación, pérdida de equipos pequeños de trabajo, etc.

1. LA SATISFACCIÓN EN EL TRABAJO DE LOS PROFESIONALES DE LA EDUCACIÓN

En la última década se han incrementado considerablemente las investigaciones sobre los roles, comportamientos y sentimientos de los profesionales de la educación. Así junto a aspectos como la formación del profesorado, han surgido estudios sobre la salud de los maestros, sobre los conflictos y malestar de los educadores, el estrés de los enseñantes, la satisfacción e insatisfacción profesional de los docentes, directivos de los Centros, Orientadores, Inspectores de Educación, etc.

Los niveles de satisfacción e insatisfacción laboral de los profesionales de la educación no son idénticos en los diferentes roles ocupacionales, ni tampoco en las diversas etapas y ciclos educativos. Incluso varían en una función profesional análoga y en el mismo nivel educativo, según sea por ejemplo, en el comienzo del ejercicio profesional o cuando se llevan varios años en el mismo.

Entre las variables que influyen en la satisfacción profesional de acuerdo con los estudios al respecto se pueden mencionar las siguientes: las condiciones laborales y el salario; la edad; el liderazgo participativo; la autonomía y variación en el rol profesional; el tamaño de la organización sociolaboral (inversamente proporcional a la satisfacción laboral); el nivel de formación y cualificación profesional; la aceptación en el grupo de trabajo.

En nuestro contexto las investigaciones sobre los niveles de satisfacción/insatisfacción de los educadores, con este término u otros como la angustia de los enseñantes, el conflicto de los profesores, la salud mental de los maestros, el estrés de los docentes, el malestar de los educadores, etc., es de interés creciente sobre todo en la última década. Como exponente de esta inquietud se pueden mencionar los estudios de Elejabeitia (1983), Esteve (1984, 1987, 1988), A. Martínez (1984), Vera (1988), González y Lobato (1988), Gordillo (1988), Gómez y Serra (1989), etc. En otros contextos socioculturales es preciso destacar las investigaciones de P. Lester (1988), A. Abraham (1986), Andrisani (1978), Fraser (1983), etc.

2. DETERMINACIÓN DE LOS NIVELES DE SATISFACCIÓN DE LOS ORIENTADORES ESCOLARES EN GALICIA

De acuerdo con los aspectos descritos anteriormente, nuestro estudio contempla los siguientes objetivos:

1.— Delimitar los factores personales, laborales, sociales, formativos, pedagógicos,... que influyen en la satisfacción/insatisfacción de los Orientadores Escolares.

2.— Analizar los motivos y niveles de satisfacción de los diferentes grupos de Orientadores, sobre todo en E.G.B. y en Enseñanza Secundaria.

3.— Determinar las categorías de satisfacción profesional referentes a las condiciones laborales, relaciones con los estamentos educativos, rol del Orientador y desarrollo de la profesión orientadora.

Entre las variables a estudiar que consideramos pueden influir en el nivel de satisfacción de los Orientadores Escolares figuran: la edad, la especialización que poseen, el sexo, el nivel educativo al que pertenecen, y su situación profesional.

Como hipótesis investigadoras partimos, de acuerdo con otras investigaciones realizadas al respecto, que los niveles de satisfacción de los profesionales de la Orientación se incrementan en los siguientes supuestos:

— Mayor edad.

— Superior nivel de formación y especialización profesional.

— Estabilidad profesional.

Según los aspectos analizados, el estudio emplea para la recogida de la información un Cuestiona-

rio de 29 ítems cerrados y 1 abierto en el que se pregunta por diversos factores y causas que influyen en la satisfacción de los Orientadores Escolares.

De las posibles metodologías de investigación a utilizar para analizar los componentes de la satisfacción profesional se usó el método directo de preguntas directas y cerradas sobre las siguientes categorías: condiciones sociolaborales de la profesión, relaciones con los alumnos, profesores y directivos escolares, actividades específicas de la profesión orientadora y desarrollo profesional del Orientador Escolar.

El cuestionario se aplicó al colectivo de 200 profesionales de la Orientación Educativa que ejercen en Galicia, en los Equipos Psicopedagógicos de Apoyo a la Escuela (nivel de Educación Básica), al que contestaron el 80%, y en los Centros de Educación Secundaria (B.U.P. y F.P.) que tienen implantado Proyectos de Orientación Educativa, que respondieron en un 60%.

3. ANÁLISIS E INTERPRETACIÓN DE DATOS Y RESULTADOS

Para el análisis de los datos e información recogida calculamos la correlación de Pearson entre la variable satisfacción y las restantes variables objeto de estudio.

Los resultados para cada una de ellas por separado se muestran a continuación:

a) Sexo

De un total de 146 sujetos, 65 son varones (44.52%) y 81 mujeres (55.47%). Los coeficientes de correlación son aquí de 0.60 y 0.75 respectivamente, revelando que como grupo las mujeres presentan un mayor grado de satisfacción en el trabajo.

b) Edad

La variable edad aparece agrupada en 3 categorías. En la primera (menos de 35 años) en la cual se incluyen 15 sujetos (10.27%) encontramos una correlación de 0.23. El segundo grupo de edad (35 a 50 años) que incluye un total de 118 sujetos (80.82%) alcanza una correlación de 0.71. Por último, para la tercera categoría de 51 a 65 años, en la que se sitúan un pequeño número de sujetos, concretamente 13 (8.90%) la correlación con satisfacción es de 0.55.

Estos hallazgos revelan que la satisfacción varía en los diferentes niveles o grupos de edad, siendo más elevada en el grupo intermedio y más baja en el grupo de menor edad (menos de 35 años).

c) Puesto Profesional

La variable puesto profesional hace referencia al cargo profesional que como Orientador se ejerce en la institución. En nuestro estudio, un total de 32 sujetos ejercen el rol de logopedas en E.G.B. y 114 ejercen el puesto de pedagogo y/o psicólogo en E.G.B. y E.E.M.M. La correlación con la satisfacción en el trabajo es de 0.76 y 0.67 respectivamente, sugiriéndonos que el grupo de logopedas está más satisfecho en su cargo que sus compañeros.

d) Nivel Educativo

Comprobamos también la relación entre el nivel educativo (E.G.B. y E.E.M.M.) y la satisfacción en el trabajo.

Para el nivel de E.G.B. (Pedagogos y/o Psicólogos) la correlación alcanzada fue de 0.71 y en el nivel de E.E.M.M., el cual incluye Bachillerato y Formación Profesional, la correlación fue de 0.61, siendo por lo tanto superior el grado de satisfacción que manifiestan los profesionales de la Orientación que trabajan en los niveles de Educación Básica.

4. CONCLUSIONES

Las conclusiones de más relieve son las siguientes:

1ª. Los niveles de satisfacción profesional son mayores entre los Orientadores y entre el grupo intermedio de edad (35 a 50 años), coincidiendo con otros estudios al respecto (Zubieta y Susinos, 1992).

2ª. Los Logopedas de los EPSA presentan mejores índices de satisfacción que sus compañeros Orientadores (Pedagogos y Psicólogos).

3ª. Los orientadores de los EPSA globalmente están más satisfechos que los profesionales de Orientación que ejercen en Centros de Educación Secundaria, lo que también es coincidente con otras investigaciones relacionadas (Johnson y Holdeway, 1990).

En síntesis los perfiles de mayor satisfacción laboral de los Orientadores escolares corresponden a los Orientadores respecto a sus compañeros, a los Logopedas Escolares que ejercen en E.G.B. respecto a Enseñanzas Medias y al grupo intermedio de edad. Los más insatisfechos son los Orientadores varones que desempeñan sus funciones en Educación Secundaria y de menor edad (menores de 35 años). Este perfil corresponde con profesionales de la Orientación con mayores expectativas y exigencias profesionales en relación con el rol académico-laboral percibido.

BIBLIOGRAFÍA

- ABRAHAM, A. y otros (1986): *El enseñante también es persona*. Barcelona: Gedisa.
- ANDRISANI, P. (1978): *Work attitudes and labor market experience*. Nueva York: Praeger Pub.
- BRUNET, L. (1987): *El clima de trabajo en las organizaciones*. México: Trillas.
- DAVIS, K. y NEWSTROUM, J. (1990): *El comportamiento humano en el trabajo*. México: Mc Graw-Hill.
- ELEJABEÍTIA, C. y otros (1983): *El maestro. Análisis de las escuelas de verano*. Madrid: EDE.
- ESTEVE, J. (1984): *Profesores en conflicto*. Madrid: Narcea.
- (1987): *El malestar docente*. Barcelona: Laia.
- (1988): «El estrés de los profesores: Propuestas de intervención para su control». En VILLA, A.: *Perspectivas y problemas de la función docente*. Madrid: Narcea/II Congreso Mundial Vasco.
- FRASER, T. (1983): *Human stress work and job satisfaction: A critical approach*. Ginebra: Oficina Internacional del Trabajo.
- GÓMEZ, L. y SERRA, E. (1989): «Sobre la salud mental de los profesores». *Cuadernos de Pedagogía*, 175, 60-65.
- GONZÁLEZ, J. y LOBATO, Mª (1988): «El malestar de los enseñantes. Problemas de salud en el País Vasco». En Villa, A.: *Perspectivas y problemas de la función docente*. Madrid: Narcea/II Congreso Mundial Vasco.
- GORDILLO, Mª (1988): «La satisfacción profesional del profesor. Consecuencias para la Orientación Educativa». En VILLA, A.: *Perspectivas y problemas de la función docente*. Madrid: Narcea/II Congreso Mundial Vasco.
- LESTER, P. (1988): *Teacher job satisfaction*. Nueva York: Garland.

- MARTÍNEZ, A. (1984): «El perfeccionamiento de la función didáctica como vía de disminución de tensiones en el docente». En ESTEVE, J.: *Profesores en conflicto*. Madrid: Narcea.
- SOBRADO, L. y GONZÁLEZ, A. (1991): «Exploración y tratamiento de conflictos en los Centros Docentes». *Educadores*, XXXIII, 160.
- VILLAR, L. (1990): *El profesor como profesional: Formación y desarrollo personal*. Granada: Servicio de Publicaciones de la Universidad.

VALORACIÓN DEL PERFIL FORMATIVO-PROFESIONAL Y DEL ACCESO DEL LOGOPEDA DE LOS EQUIPOS PSICOPEDAGÓGICOS DE SECTOR ESCOLAR

Luis Sobrado Fernández, Ana M^a Porto Castro y Leonor Filgueira Domínguez
Universidad de Santiago

INTRODUCCIÓN

Entre los rasgos característicos que distinguen a un buen profesional figuran en primer lugar la *competencia*, es decir, el «saber como», y por ello una persona competente en una profesión posee una formación completa y amplia en la misma.

El profesionalismo debe basarse en una fundamentación sólida de formación y actitudes que guían a la persona en el trabajo que desempeña. Generalmente el nivel de profesionalización con el que se realiza una ocupación o trabajo está vinculado con la capacitación y la eficacia en su desarrollo y con una proyección moral.

De acuerdo con lo expresado anteriormente se entiende la *profesionalización* como el sistema por el que una persona que ejerce una ocupación o empleo (en función de una competencia especial, de la calidad de su actividad y de los beneficios que reporta a la sociedad) posee el derecho a efectuar una modalidad de trabajo determinado (Clamp 1989).

La profesión de Logopeda Escolar de un Equipo de Orientación Educativa de Sector es una ocupación que exige un conocimiento especializado y unas competencias y habilidades específicas que implican una formación sistemática y regular tanto inicial como permanente.

1. PROFESIONALIZACIÓN DEL LOGOPEDA DEL EQUIPO DE ORIENTACIÓN PSICOPEDAGÓGICA DE ZONA ESCOLAR

La identidad profesional del Logopeda Escolar se deriva de su formación como especialista en Audición y lenguaje, basada sobre todo en las Ciencias de la Educación, y de una preparación concreta en las habilidades psicopedagógicas adaptadas al contexto educativo y sociolingüístico en el que realiza su labor.

Los profesionales de la Audición y Lenguaje, de Equipos Psicopedagógicos de distrito escolar, desempeñan roles distintos en los diferentes niveles educativos (infantil, primaria, educación especial, etc...) de acuerdo con las diferentes etapas de desarrollo evolutivo de los niños, adolescentes, etc... Asimismo deben estimular el desenvolvimiento y ayuda a los alumnos en colaboración con otros profesionales como son los trabajadores sociales, profesores, psicólogos, pedagogos, médicos, etc...

Por otra parte los Logopedas deben facilitar a las familias un conocimiento adicional del desarrollo personal de sus hijos, y fortalecer el rol de éstas en la estimulación de un desenvolvimiento equilibrado de los niños y adolescentes.

Los profesionales de la Logopedia Escolar deben enfatizar las siguientes *intervenciones* en los distintos niveles educativos:

1ª Promover en los escolares aspectos del desenvolvimiento humano y lingüístico, a través de la estructuración de las experiencias orientadoras por medio de grupos, o bien integrándolas en el Currículum Escolar, por ejemplo en el área de Lengua. Se les debe facilitar asimismo orientación personal cuando las diversas necesidades lingüístico-educativas exigen una atención individualizada.

2ª Asesoramiento al profesorado para que incremente sus habilidades de comunicación, mejore sus destrezas de intervención tutorial con los alumnos y sus familias, y se sensibilice en la conveniencia de integrar las necesidades educativas y de desarrollo lingüístico de éstos en el Currículum Educativo.

2. NUEVOS PERFILES PROFESIONALES DEL LOGOPEDA ESCOLAR COMO ORIENTADOR

La educación precisa actualmente cambios importantes y ello conlleva la necesidad de modificaciones concomitantes en la Orientación Educativa y Profesional y consecuentemente en el rol profesional de los Logopedas Orientadores.

Respecto a su perfil parece existir un consenso en que debe ser un especialista en el conocimiento del desarrollo lingüístico del escolar, que debe integrarlo en los currículos escolares, que debe de elaborar y aplicar programas de prevención primaria en el área lingüística y que a través de éstos se organicen ambientes saludables de aprendizaje y mejores climas escolares (Thomas, 1989).

La *consulta* parece un sistema adecuado de estrategia personal de los Logopedas-Orientadores considerando su trabajo enfocado a una amplia variedad de destinatarios: Alumnos, familias, profesores, equipos directivos de los centros, etc. El procedimiento de intervención a través de la consulta parece efectivo, tanto para los objetivos de la prevención como para la intervención orientadora, y existen estudios que la apoyan como sistema que incrementa el éxito escolar (Thomson, 1987).

Por otra parte como señalan Childers y Podenski (1987) la orientación grupal suele ser un método apropiado para emplearse por los profesionales de la Orientación.

3. FORMACIÓN DE LOGOPEDAS-ORIENTADORES PARA EL CAMBIO SOCIO-EDUCATIVO

¿Cómo se debe de preparar a los Logopedas Escolares para que sean capaces de moverse en un entorno escolar con nuevas perspectivas, paradigmas y destrezas de cambio?

La Carta de Educación de Harvard de 1988 sugiere cuatro estrategias para ayudar a los Orientadores Escolares a que respondan a los nuevos roles y funciones, y son:

1ª Reforzar las vinculaciones entre la Orientación y el Currículum Educativo.

2ª Incrementar el número de profesionales de la Orientación, incluidos los Logopedas, sobre todo en los niveles de educación infantil y primaria, y comenzar a tratar de un modo preventivo y temporal los problemas de desarrollo personal, social y lingüístico de los niños.

3ª Establecer mayores vínculos y servicios con las familias y más enlaces y coordinación entre el Centro Escolar y los padres, a través de programas cooperativos de educación.

4ª Atender más directamente a los problemas de la Comunidad Educativa y social a través de sistemas de ayuda y respuesta a las necesidades culturales y sociales de la misma.

Los Logopedas-Orientadores de sector escolar deben poseer una amplia formación en estas habili-

dades que reclaman la existencia de especialistas en el desarrollo de programas de intervención orientadora en el área lingüística, en sistemas de consulta para alumnos, profesores y padres, en estrategias de prevención y estimulación temprana, en asesoramiento multicultural y multilingüe, en intervención familiar, en orientación grupal y comunitaria, etc.

Los programas de formación de Orientadores-Logopedas Escolares, deben abordar estas nuevas funciones y destrezas con precisión y rigor, sin esperar a que se aprendan en el propio ejercicio profesional.

4. METODOLOGÍA

El presente estudio sobre el perfil del Logopeda de los Equipos Psicopedagógicos de Sector Escolar (EPSA) tiene como objetivo investigar la **formación** y el **acceso** a dicho puesto en la Comunidad Autónoma Gallega, y a partir de aquí, esbozar el perfil formativo-profesional de estos Orientadores.

Las variables consideradas en el estudio atienden a los siguientes aspectos:

A) Formación inicial y continua.

B) Acceso al puesto profesional.

Para la recogida de información se les aplicó un cuestionario «ad hoc» de 63 ítems en total, contestado por 33 miembros de un total de 38 (87%).

5. ANÁLISIS E INTERPRETACIÓN DE LOS DATOS

En cuanto a la *formación*, el estudio y análisis de los datos recogidos en el mismo nos indica que la casi totalidad de los especialistas en Audición y Lenguaje (90,9%) opinan que ésta debe de ser fundamentalmente teórico práctica y la institución que debe impartirla debe de ser la Universidad principalmente (39,4%) seguida de las Asociaciones profesionales (21,2) y de la Administración Educativa (18,2%).

Por lo que se refiere a la duración de los programas, para una formación adecuada, casi la mitad de estos profesionales (45,5%) piensa que los cursos básicos de formación inicial deberían de contar con más de 500 horas.

En lo que atañe a la formación continua el 33,3% de los sujetos piensa que la oferta actual es insuficiente y un 27,3% la considera muy escasa. Esta debería realizarse tanto dentro como fuera de la sede de los EPSA, es decir, en una situación mixta y valoran preferentemente los **Seminarios permanentes**, como la modalidad más adecuada para dicha formación, seguida de la **autoformación** a través de asistencia a cursos y lectura de revistas, y de la participación en jornadas y congresos.

En relación a la *selección* para la función profesional de Logopeda de los EPSA, entre las posibles modalidades actuales, se da preferencia al perfil de Maestro Licenciado en Pedagogía y/o Psicología y especialista en Audición y Lenguaje.

Los datos también nos revelan que un amplio porcentaje de la muestra, concretamente el 63,6%, considera el concurso de **méritos** como la modalidad de selección más idónea, si bien, un 24,2% opina que sólo deberían de ser seleccionados los Logopedas que ejercen en los Centros docentes a través de un concurso restringido.

6. CONCLUSIONES

1ª) La formación inicial de los Logopedas de los EPSA debe ser teórico-práctica, impartida por la

Universidad en colaboración con Asociaciones Profesionales y con la Administración Educativa, y de amplia duración (más de 500 h.).

2ª) La modalidad de formación continua por la que abogan es prioritariamente la de Seminario Permanente. Piensan que la oferta actual es escasa.

3ª) Respecto al sistema de acceso a los Servicios de Orientación de sector, consideran que debe de ser a través de un **concurso de méritos** (modalidad actual) restringiéndose éste en el futuro a los Logopedas de los Centros docentes (en la actualidad casi inexistentes), y muestran preferencia por el perfil del Maestro Licenciado en Pedagogía y/o Psicología y especialista en Audición y Lenguaje.

BIBLIOGRAFÍA

- CARNEGIE FORUM ON EDUCATION (1986): *A nation prepared*. Nueva York.
- CHILDERS, J. H. y PODEMSKI, R. S. (1987): «Implementing group work in secondary schools. Leadership steeps for principals». *NASSP Bulletin*, 71, 83-88.
- CLAMP, P. C. (1989): «Professionalism: A state of mind». *Education Canada*, 29, 4, 12-15.
- ECHEVERRÍA, B. (1993): ¿Nuevas cualificaciones del Orientador o una nueva forma de interpretarla profesión? Tenerife: V Seminario Iberoamericano de Orientación (13-16 de mayo de 1992).
- MEC. (1990): *La Orientación Educativa y la intervención Psicopedagógica*. Madrid.
- REPETTO, E. (1993): Perspectivas de la formación de Orientadores. El problema de la capacitación en una dimensión europea. Tenerife: V Seminario Iberoamericano de Orientación (13-16 de mayo de 1992).
- SOBRADO, L. (1993): *Intervención Psicopedagógica y Orientación Educativa*. Barcelona. P.P.U.: 3ª Edición.
- THOMAS, M. D (1989): «The Counselor in effective schools». *The School Counselor*, 36, 249-252.
- THOMPSON, R. A. (1987): «Creating instructional and counseling partnerships to improve the academic performance of underachievers». *The School Counselor*. 34, 289-296.
- WATTS, A. G. (1992): Perfiles profesionales de los consejeros de Orientación en la CEE. Informe de síntesis. CEDEFOP: Berlín.

LA LABOR ORIENTADORA DEL PROFESOR UNIVERSITARIO DESDE LA PERSPECTIVA DEL ALUMNADO Y DEL PROFESORADO

Feli Arbizu Bakaikoa
I.C.E. de la Universidad del País Vasco

1. INTRODUCCIÓN

El conocimiento y la actitud del profesorado y del alumnado hacia las tareas que constituyen la docencia universitaria, representa uno de los puntos clave de todo proceso de mejora docente (Kyriacou, 1991). Nosotros pretendemos, a través de esta comunicación, acercarnos a una parte muy importante de la docencia universitaria: la Función Orientadora de la misma.

La investigación sobre el tema de la orientación universitaria es un campo, que en ámbito de la universidad española no ha cubierto demasiadas etapas, si bien el interés por el mismo ha ido creciendo a lo largo de estos últimos años.

A través de esta comunicación pretendemos incitar al debate acerca de una serie de aspectos que nos lleven a considerar que acciones orientadoras deberían ser desarrolladas desde la docencia universitaria. Así pues vamos a presentar los objetivos marcados para esta comunicación:

— Analizar la importancia otorgada por el alumnado y la responsabilidad asumida por el profesorado con respecto a las tareas y subfunciones que constituyen la Función Orientadora universitaria.

— Analizar el nivel de percepción de profesores y alumnos sobre la realización por el profesorado de cada una de las tareas y subfunciones de la Función Orientadora.

— Analizar el nivel de necesidad manifestado por profesores y alumnos en torno a las tareas y subfunciones de la Función Orientadora.

2. LA FUNCIÓN ORIENTADORA EN EL ÁMBITO DE LA UNIVERSIDAD

La docencia universitaria constituye el desarrollo de acciones que tienen como objetivo la formación del alumno universitario. En este sentido es preciso subrayar el carácter heterogéneo pero interactivo del quehacer docente. La docencia no se limita a la enseñanza que realiza un profesor concreto en el aula, sino que supone la realización de tareas de alta complejidad que, a veces, corresponde desarrollar no sólo a un profesor sino a un conjunto de ellos (De Miguel, 1991). La identificación de esas acciones, supone en sí el sustrato de la definición del modelo de perfil profesional docente.

Entre toda la serie de acciones a desarrollar desde la docencia universitaria están aquellas diferen-

ciadas y que constituyen la vertiente orientadora de la docencia, que nosotros denominamos Función Orientadora. Cualquier análisis de los elementos constituyentes de una Función y en este caso de la Función Orientadora supone la determinación del conjunto de elementos que la componen y las relaciones recíprocas entre ellas en orden a la finalidad para la que se organiza dicha función. Este análisis de funciones supone analizar lo que debe hacerse y establecer el orden de las mismas en una estructura.

Las acciones orientadoras a desarrollar desde la vertiente docente universitaria pueden ser objeto de largos debates. El contenido de las tareas que deben configurar la Función Orientadora, creemos que es una puerta abierta, no solamente en relación a lo que señalan los especialistas, sino sobre todo en razón de un análisis de necesidades contextualizado.

La realidad de cada centro nos informa sobre las necesidades que sobre orientación tienen los alumnos de cada centro (Díaz Allué, 1989). Del mismo modo el perfil profesional de cada especialidad nos obliga a considerar aquellas acciones orientadoras necesarias para la adecuada formación de los alumnos que acceden a dicha especialidad. Finalmente será necesario considerar las condiciones en la que se desarrolla dicha acción orientadora; la preparación del profesorado para desarrollar dichas acciones, los recursos disponibles, la distribución de tiempos de dedicación del profesorado, etc.

En razón de ello, con nuestra propuesta no pretendemos señalar cuales deben de ser los elementos de acción que configuran la Función Orientadora universitaria. Sin embargo dado que pretendemos conocer el punto de vista del alumnado y del profesorado acerca de dicha función, hemos presentado a la comunidad universitaria un modelo para que pueda posicionarse sobre él. El modelo ofrecido como referencia tiene en cuenta acciones o tareas que supuestamente un profesor universitario debe realizar en el ámbito de la orientación y que se identifican con las tres áreas, tradicionalmente consideradas, de la orientación educativa: la orientación personal, la académica y la profesional. La formulación de dichas tareas responde a las aportaciones de diversos autores acerca de acciones que un profesor debe realizar en el ámbito de la orientación (Villa, 1985, Centra, 1980, Forero, 1984, Mateo, 1990). De este modo las tareas orientadoras a realizar dentro de la acción docente universitaria y que han sido objeto de encuesta son doce. Las cuatro primeras forman parte de la subfunción denominada: Orientación Personal, las tareas 5, 6, 7, 8 y 9 forman parte de la Orientación Académica y, finalmente las tres últimas forman parte de la subfunción denominada Orientación Profesional.

Subfunción: Orientación Personal

- 1— Orientar al alumno en cuestiones de orden personal.
- 2— Fomentar en el alumno valores, actitudes positivas.
- 3— Buscar soluciones para algunas de las situaciones personales del alumno.
- 4— Orientar al alumno en valores morales y éticos.

Subfunción: Orientación Académica

- 5— Comunicar al alumno sus facultades y dificultades y orientarle para que supere las dificultades y potencie sus facultades
- 6— Orientar al alumno en las materias, métodos de estudio, técnicas...
- 7— Aplicar métodos de recuperación al alumno con dificultades
- 8— Desarrollar en el alumno las habilidades necesarias de la profesión para la que se prepara.

Subfunción: Orientación Profesional

- 9— Informar al alumno de cómo funciona la universidad: estructura, biblioteca...
- 10— Informar al alumno sobre salidas profesionales de su carrera.

11— Mostrar las habilidades, técnicas, etc. que necesita desarrollar el alumno para ejercer la profesión para la que se prepara.

12— Cotactar al alumno con el mundo laboral, haciéndole participar en proyectos con empresas...

3. PRECISIONES METODOLÓGICAS

Debemos comenzar señalando que los resultados presentados en esta comunicación forman parte de una investigación cuyo objetivo era analizar y definir la Función Docente universitaria en el contexto de la Universidad del País Vasco (Arbizu, 1992). Para ello se construyó un cuestionario constituido por 130 ítems que hacían referencia a otras tantas tareas que supuestamente debe desarrollar el profesor universitario. En esta comunicación sólo presentamos el análisis específico de una parte del cuestionario, en concreto la referida a la Función Orientadora, que consta de 12 ítems.

La muestra utilizada está constituida por un colectivo de 45 profesores y 727 alumnos de 5 centros de la Universidad del País Vasco (tres Facultades y dos Escuelas Universitarias).

A los alumnos se les preguntaba en primera instancia, sobre el grado de importancia que concedía a cada una de las tareas que se señalaban en el cuestionario. El encuestado debía responder en torno a una escala de 1 a 10, donde el 1 indicaba que no consideraba importante dicha tarea y el 10 indicaba que la consideraba de total importancia. Después de ello se les preguntaba sobre el grado en que consideraban que sus profesores realizaban dichas tareas. En este caso, también, la escala de respuestas iba de 1 a 10, siendo el 10 el máximo grado de realización y el 1 el grado nulo de realización.

A los profesores se les preguntaba, en primer lugar sobre el grado en que consideraban que cada una de las tareas presentadas formaban parte de su responsabilidad docente. La escala de respuesta iba, igualmente de 1 a 10, siendo 10 el grado máximo de responsabilidad y el 1 indicaba que no había responsabilidad alguna. En segundo lugar, se les preguntaba por el grado en que realizaban cada una de las tareas. En la escala de respuestas el 1 indicaba que no realizaban dicha tarea y el 10 que la realizaban en máximo grado.

4. OPERATIVIZACIÓN DEL CONCEPTO DE NECESIDAD

Para analizar los grados de percepción de la importancia otorgada por el alumnado y la responsabilidad asumida por el profesorado respecto a las tareas presentadas, así como los niveles de realización con respecto a las mismas, se hizo uso de las puntuaciones directas otorgadas por los encuestados. Sin embargo para lograr el tercer objetivo, analizar los niveles de necesidad fue necesario operativizar el concepto de necesidad.

Para ello definimos el concepto de necesidad como la discrepancia existente entre la situación dada y la deseada del desarrollo educacional (Tejedor, 1990). Una necesidad viene constituida por la diferencia o discrepancia entre la forma en cómo son las cosas realmente y la forma en que deberían ser (exigencia), podrían ser (posibilidades reales) o nos gustaría que fueran (necesidades individualizadas) (Zabalza, 1991).

Una necesidad implica un juicio de valor que es emitido por un grupo social. Por ello somos conscientes de que las circunstancias y condiciones del grupo que se toma como elemento de referencia de la detección de necesidades influyen en el concepto de necesidad manifestado por el grupo. En nuestro caso los grupos de referencia elegidos eran los profesores y alumnos. Sabemos que la realidad de cada uno de los colectivos era diferente, pero comparten una misma acción, la docente. Los profesores, como ejecutores de la misma y el grupo de alumnos como elementos a los que se destina la acción docente.

Teniendo en cuenta la definición de necesidad adoptada, la operativización del concepto de

necesidad, en el caso del alumnado, se hizo restando al grado de importancia otorgado a cada tarea docente, el grado en que consideran que sus profesores realizan cada una de ellas.

IMPORTANCIA — REALIZACIÓN = NECESIDAD

En el caso del profesorado, la operativización del concepto de necesidad, se hizo restando al grado de responsabilidad asumida, el grado de realización considerado.

RESPONSABILIDAD — REALIZACIÓN = NECESIDAD

Hechas las apreciaciones metodológicas, así como la explicación de la operativización del concepto de necesidad, procederemos a presentar el análisis de resultados en razón de los objetivos que nos hemos marcado.

5. GRADOS DE IMPORTANCIA OTORGADOS POR EL ALUMNADO Y GRADOS DE RESPONSABILIDAD ASUMIDOS POR EL PROFESORADO CON RESPECTO A LAS SUBFUNCIONES Y TAREAS DE LA FUNCIÓN ORIENTADORA

Como se puede apreciar en el Gráfico 1, los perfiles de alumnado y profesorado son paralelos en cuanto a las tres subfunciones que constituyen la Función Orientadora del profesor universitario.

Se puede apreciar que la importancia otorgada por el alumnado, tanto a las subfunciones como a las tareas, es mayor que la responsabilidad asumida por el profesorado respecto a ellas.

Hay coincidencia, casi total, respecto a las tareas de la subfunción de Orientación Personal. Como se puede apreciar las tareas de esta subfunción son las que obtienen puntuaciones más bajas, excepto la tarea nº 2: «Fomentar en los alumnos valores e ideas positivas», que obtiene puntuaciones apreciables tanto en cuanto a la importancia otorgada por el alumnado como por la responsabilidad asumida por el profesorado. Sin embargo el resto de tareas obtienen puntuaciones por debajo del 5,5, límite, establecido por nosotros, por debajo del cual podríamos considerar que los grados de importancia o de responsabilidad no son aceptables.

En cuanto a las Subfunción de Orientación Académica, podemos observar que el alumnado otorga niveles de importancia altos, siendo la tarea nº 8. «Desarrollar en el alumnado las habilidades necesarias para la profesión para la que se preparan» la que tiene una puntuación más alta (8,8). En cuanto al profesorado, observamos que si bien asume niveles de responsabilidad, sin embargo éstos están por debajo de los niveles de importancia otorgados por alumnos.

Respecto a las Subfunción de Orientación Profesional, observamos que para los alumnos es la que obtiene niveles de importancia más altos dentro de la Función Orientadora. Las tres tareas que la componen obtienen puntuaciones muy altas y similares.

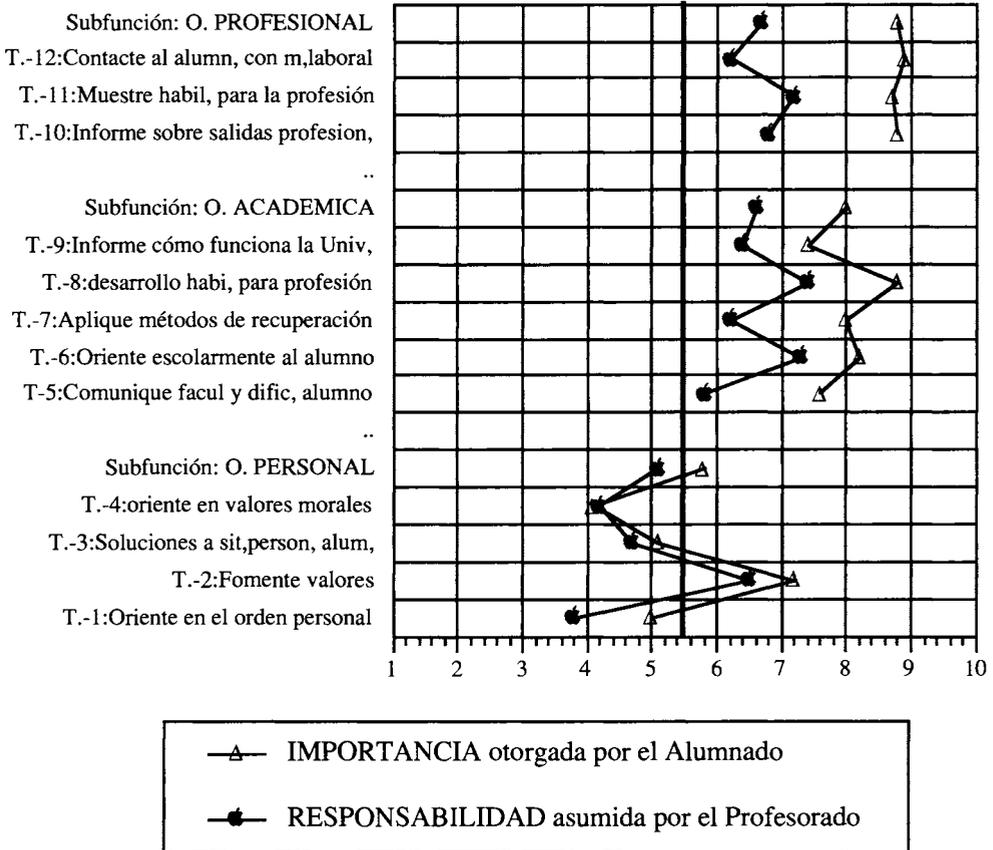
6. PERCEPCIÓN DE LOS GRADOS DE REALIZACIÓN DE LAS SUBFUNCIONES Y TAREAS DE LA FUNCIÓN ORIENTADORA

Si observamos el gráfico 2, podemos comprobar que las percepciones del profesorado y del alumnado marcan perfiles bastante paralelos. En todos los casos el alumnado considera que el nivel de realización por parte del profesorado es apreciablemente más bajo que el considerado por los propios profesores.

Según el alumnado el nivel de realización de las diferentes tareas y subfunciones por parte de sus profesores es muy baja, pues las puntuaciones se encuentran entre el 2 y el 3, lo que indica que para el

GRÁFICO 1

GRADOS DE IMPORTANCIA OTORGADOS POR EL ALUMNADO Y GRADOS DE RESPONSABILIDAD ASUMIDOS POR EL PROFESORADO CON RESPECTO A LAS SUBFUNCIONES Y TAREAS DE LA FUNCIÓN ORIENTADORA

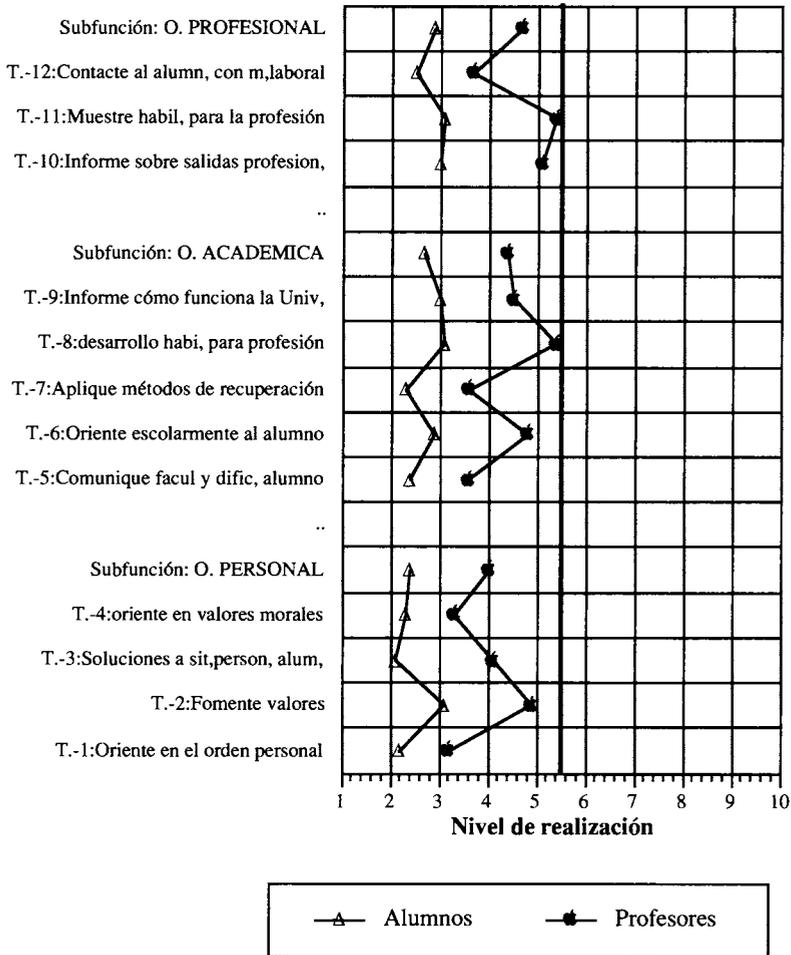


alumnado el nivel de realización de los profesores con respecto a todas las tareas de la Función Orientadora es bajísima, casi nula. (recuérdese que el 1 indica un nivel de realización de 0).

En cuanto al colectivo de profesores puede apreciarse un perfil con puntuaciones entre el 3 y el 4,4. Ello indica que el profesorado reconoce que la realización de estas tareas es muy baja. Podemos observar que todas las puntuaciones se sitúan en la mitad izquierda (Recuérdese que hemos establecido en el 5,5 el punto medio, de modo que aquellas puntuaciones inferiores a ese número son consideradas como niveles de realización no aceptables).

En cuanto a la diferenciación de los niveles de realización con respecto a las diversas subfunciones, vemos que según el alumnado son muy similares.

GRÁFICO 2
 GRADOS DE REALIZACIÓN PERCIBIDOS POR EL ALUMNADO Y PROFESORADO CON
 RESPECTO A LAS SUBFUNCIONES Y TAREAS DE LA FUNCIÓN ORIENTADORA



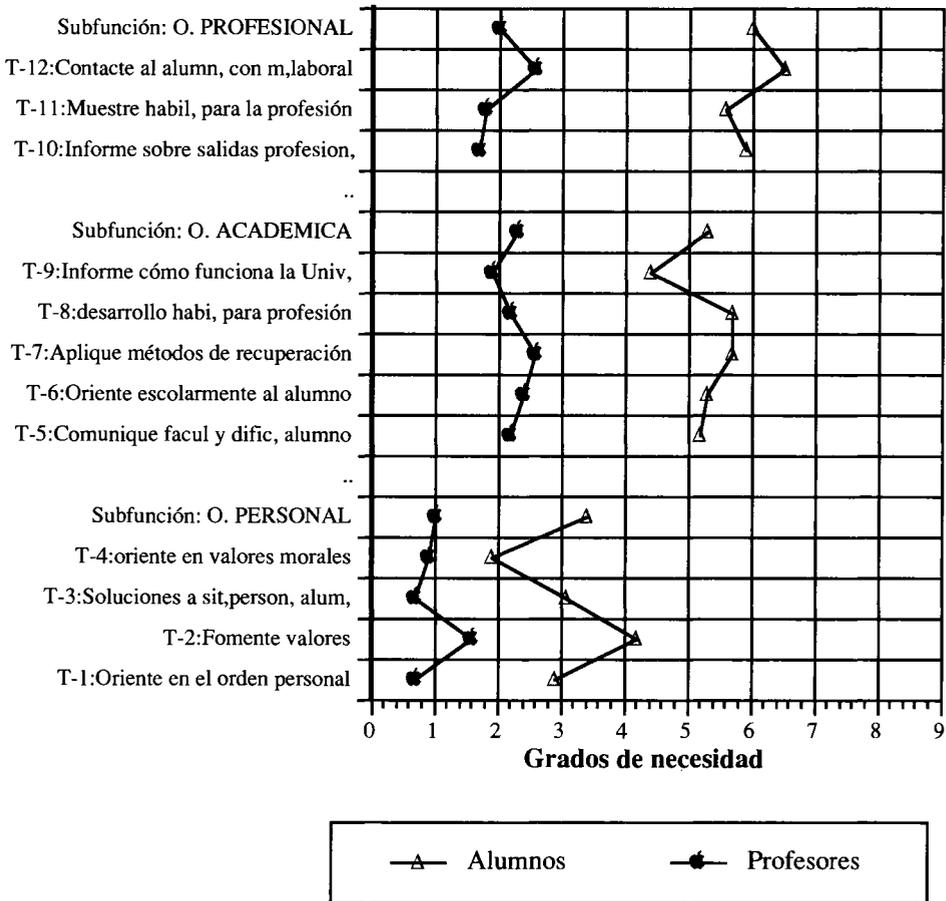
Para el profesorado hay una mayor diferencia entre los niveles de realización de las tareas. Sin embargo esta diferencia casi desaparece cuando observamos las puntuaciones medias de las subfunciones. De este modo se puede observar que dentro de la subfunción de Orientación Profesional, el profesorado reconoce que la tarea que menos realizan es la nº 12: «Contactar al alumno con el mundo laboral, haciéndole participar en proyectos con empresas...». En cuanto a las tareas de la subfunción de Orientación Personal, la tarea nº 1: «Orientar al alumno en cuestiones de orden personal» es la que consideran que realizan en menor medida. En este sentido debe recordarse que es la tarea que obtiene el menor nivel de responsabilidad por parte de los profesores.

7. PERCEPCIÓN DE LOS GRADOS DE NECESIDAD DE LAS SUBFUNCIONES Y TAREAS DE LA FUNCIÓN ORIENTADORA

Pasando al tercer gráfico que presenta los niveles de necesidad de los dos colectivos encuestados, podemos observar que existe una mayor distanciamiento entre las puntuaciones de los dos colectivos y que éstas varían respecto a cada una de las tres subfunciones.

Como primer comentario de este apartado conviene resaltar que los grados de necesidad de los alumnos son apreciablemente mayores que los de los profesores.

GRÁFICO 3
GRADOS DE NECESIDAD PERCIBIDOS POR EL PROFESORADO Y ALUMNADO CON RESPECTO A LAS SUBFUNCIONES Y TAREAS DE LA FUNCIÓN ORIENTADORA



En cuanto al alumnado comprobamos que tanto la subfunción de Orientación Académica, pero sobre todo la de Orientación Profesional obtiene niveles de necesidad bastante altos, destaca entre todas, la tarea nº 12: «Contacte al alumno con el mundo laboral...».

En cuanto a la subfunción de Orientación Personal, comprobamos que para el alumnado los niveles de necesidad son medio bajos, observándose unas puntuaciones muy dispersas en las tareas que componen dicha subfunción. Entre dichas tareas la que parece obtener un grado de necesidad más alto es la nº 2: «Fomentar en los alumnos valores e ideas positivas»

En cuanto al profesorado como ya hemos señalado los niveles de necesidad parecen ser bastante menores, siendo la subfunción de Orientación Personal la que obtiene puntuaciones más bajas.

8. CONCLUSIONES

Las posibilidades de analizar la Función Orientadora universitaria son amplias y diversas. Nosotros no hemos hecho sino abrir una vía para la continuación de una línea de reflexión y profundización. Sin embargo, nos parece oportuno terminar esta comunicación, destacando una serie de consideraciones que hemos deducido de nuestro análisis.

— En general, la diferencia existente en las percepciones con respecto a los componentes de la Función Orientadora, nos indica que existe un desfase que es preciso analizar en profundidad.

— Se constata que los niveles de realización de las tareas que constituyen la Función Orientadora universitaria por parte del profesorado, están muy lejos de alcanzar los niveles óptimos de desarrollo.

— Se confirma, ampliamente, la diferencia de percepción entre profesorado y alumnado en cuanto a las necesidades de orientación (Calkins & Wakeford, 1983).

— El alumno universitario demanda de un modo urgente una serie de acciones orientadoras. Estas acciones, fundamentalmente deben estar centradas en acciones de orientación académica y profesional. El alumno universitario precisa algo más que contenidos a aprender y prácticas a realizar, necesita ayuda para autocomprenderse, para adquirir conciencia de la realidad y penetrar en las perspectivas de futuro, sin que ello se traduzca en una «domesticación»; se hace imprescindible desarrollar procesos de orientación y asesoramiento (Echeverría, 1988)

— Si bien son preocupantes los bajos niveles de realización de las tareas orientadoras, hecho reconocido por el propio profesorado, resulta esperanzador que el profesorado acepte que dichas tareas forman parte de sus competencias docentes, lo que permite augurar que con una adecuada preparación y una asignación de medios el profesorado podrá responder a las necesidades manifestadas por el alumnado en el ámbito de la orientación universitaria.

— En nuestras instituciones la actividad orientadora suele ser muy reducida y asistemática. Normalmente, se reduce a actividades ligadas a la disciplina como son la búsqueda bibliográfica, la orientación de trabajos prácticos, etc. (Pérez Juste, 1985). Pero actividades más propiamente orientadoras como la formación en técnicas de estudio, la elección de asignaturas optativas, la orientación de tipo profesional u otras, no suelen ser muy habituales en nuestros centros. Por ello la acción orientadora del profesorado universitario se hace evidente y debe traducirse en la ayuda al alumnado en cuanto a: elección de asignaturas, ayuda en la toma de decisiones o en la realización de trabajos, en la orientación profesional y en la tarea de contactar al alumno con el mundo del trabajo. En este sentido el sistema orientador universitario inglés, bien nos puede servir de guía, si bien es necesario reconocer que no es fácilmente trasvasable a nuestro contexto (Gordillo, 1985). Sin embargo, no impide que ciertas acciones puedan y deban ser desarrolladas en nuestras universidades.

— Por otra parte, es cierto que el profesor no puede abarcar toda la serie de acciones, algunas de ellas muy especializadas, que requiera una buena orientación. Sin embargo, puede desarrollar algunas tareas orientadoras que, si se ven complementadas con un buen servicio de orientación universitaria,

conseguirán satisfacer la gran demanda de orientación que los alumnos reclaman. No obstante, cualquier servicio de orientación universitaria poco podrá desarrollar si no cuenta con la acción integradora de la orientación de cada profesor (Escotet, 1992).

— Así pues, consideramos la necesidad de que la Universidad establezca un servicio de orientación al alumnado. No obstante, esto no significa que el profesor quede exento de dichas tareas, muy al contrario, debe desarrollarlas e integrarlas dentro de su acción docente. Negar la acción orientadora que el profesor debe desarrollar supone que el profesor se constituye en el elemento que ayuda al alumno a conseguir las metas. Evidentemente, esto implica un mínimo conocimiento del alumno, de su desarrollo y de cómo puede superar sus dificultades (Pérez Juste, 1985).

REFERENCIAS BIBLIOGRÁFICAS

- ARBIZU, F. (1992): *La Función Docente del profesor universitario*. Tesis Doctoral. Bilbao: Universidad del País Vasco.
- CALKINS, E. V. & WAKEFORD, R. (1983): Perception of instructors and student of instructors' roles. *Journal of Medical Education* 58 (Dec), 967-969.
- CENTRA, J. A. (1980): *Determining faculty performance*. San Francisco: Jossey Bass.
- DE MIGUEL, M. (1991): Indicadores de Calidad de la Docencia Universitaria. In *I Congreso Internacional sobre Calidad de la Educación*. Cádiz: ICE.
- DÍAZ ALLUÉ, M^a T. (1989): *La orientación universitaria ante la problemática académica y profesional del estudiante*. Madrid: Narcea.
- ECHEVERRÍA, B. (1988): Modelos de institucionalización de la orientación educativa y sus implicaciones en la calidad de la educación. In *IX Congreso Nacional de Pedagogía: La Calidad de los centros educativos*, (pp. 393-411). Alicante: Sociedad Española de Pedagogía.
- ESCOTET, M. A. (1992): *Aprender para el futuro*. Madrid: Alianza Editorial.
- FORERO, F. (1984): El programa de Magister en docencia universitaria. *Documentación Educativa (Colombia)*, 8 (20), 235-244.
- GORDILLO, M. V. (1985): La orientación en la Universidad. *Revista Española de Pedagogía*, XLIII (169-170), 435-451.
- KYRIACOU, C. (1991): Developing your teaching skills. KYRIACOU (Ed.): *Essential Teaching Skills*. Oxford: Basil Blackwell.
- MATEO, J. (1990): Función docente y demanda discente. *Revista española de pedagogía*, 186 (XL-VIII), 319-326.
- PÉREZ JUSTE, R. (1985): Programación de actividades universitarias. *Revista Española de Pedagogía*, (169-170), 391-407.
- TEJEDOR, F. J. (1990): Perspectivas metodológicas del diagnóstico y evaluación de necesidades en el ámbito educativo. In *RIE*, 8 (pp. 15-37).
- VILLA, A. (1985): *Un modelo de profesor ideal. Multidimensionalidad del modelo del profesor ideal y condiciones estructurales que lo determinan*. Madrid: MEC.
- ZABALZA, M. A. (1991): *Diseño y desarrollo curricular*. 4^a ed. Madrid: Narcea.

COMUNICACIONES AL SIMPOSIUM II

¿DIAGNÓSTICO PEDAGÓGICO O EVALUACIÓN EDUCATIVA? HACIA UNA CLARIFICACIÓN CONCEPTUAL Y TERMINOLÓGICA

Raquel-Amaya Martínez González
Universidad de Oviedo

El contenido de esta comunicación versa sobre los conceptos de Diagnóstico Pedagógico y Evaluación Educativa. El objetivo que se persigue es clarificar, en la medida de lo posible, lo que ambos términos significan, tratando de delimitar sus aspectos comunes y diferentes. Para ello se establecen dos grandes apartados. En el primero revisaremos el grado de sinonimia de estos términos en virtud de las definiciones y conceptualizaciones que diversos autores han ofrecido sobre ellos. En el segundo trataremos de delimitar algunos aspectos diferenciales que permitan avanzar hacia su clarificación conceptual.

1. ¿DIAGNÓSTICO PEDAGÓGICO O EVALUACIÓN EDUCATIVA?

Para poder establecer si existe sinonimia entre ambos términos es necesario repasar algunas conceptualizaciones y definiciones que se han emitido sobre ellos. Comenzaremos analizando las referidas a Evaluación Educativa por ser más numerosas y conocidas, y posteriormente, nos adentraremos en las relativas a Diagnóstico Pedagógico.

1.1. Evaluación educativa

La evaluación, de acuerdo con Orden Hoz (1989), es «una función básica e ineludible en el proceso de intervención educativa». Pero cuando uno se acerca a los textos específicos que tratan sobre ella para clarificar su significado, se advierte que no existe un consenso entre los especialistas del que pueda derivarse una definición unívoca de este término.

Esta diferencia de significados viene marcada, fundamentalmente, por el énfasis que se otorga a determinadas características de la evaluación (Orden Hoz, 1991), y por las metas que se asignan tanto a la actividad evaluadora como al rol del evaluador desde diferentes concepciones teóricas. Tomando este último criterio —el rol del evaluador—, Nevo (1990) sintetiza estas concepciones en tres fundamentales: (1) evaluación dirigida a la consecución de metas, (2) evaluación descriptiva no valorativa dirigida a proporcionar información para la toma de decisiones o para examinar programas sociales y

educativos, y, por último, (3) evaluación valorativa, que busca establecer el valor o mérito de aquello que se evalúa, o bien una combinación de descripción y valoración.

Walberg y Haertel (1990:xvii), definen la evaluación como «a careful, rigorous examination of an educational curriculum, program, institution, organizational variable, or policy», cuyo primer objetivo es llegar a conocer la entidad estudiada para mejorarla (evaluación formativa), juzgar los resultados previstos e imprevistos (evaluación sumativa), o ambas cosas a la vez. En cualquier caso, se realiza, según estos autores, sobre programas, prácticas o materiales que están siendo utilizados para un objetivo concreto.

Popham (1988:7) habla de «evaluación educativa sistemática», y la entiende como «a formal appraisal of the quality of educational phenomena», que se diferencia de la evaluación informal que se realiza a diario para valorar aspectos de la vida cotidiana. Entiende esa comprobación de calidad, además, en términos de valor o mérito del fenómeno que se evalúa, excluyendo de este fenómeno los aprendizajes o ejecuciones de los alumnos sometidos al proceso de enseñanza-aprendizaje, que son valorados, de acuerdo con este autor, a través de otro proceso denominado «grading».

En nuestro contexto educativo, el profesor Orden Hoz (1991) ofrece una definición bastante comprensiva. Entiende que la evaluación «es el proceso o conjunto de procesos sistemáticos de recogida, análisis e interpretación de información relevante para medir o describir cualquier faceta de la realidad educativa, y sobre esta descripción formular un juicio de valor por su comparación con un criterio o patrón, todo ello como base para tomar decisiones».

Como puede apreciarse por las definiciones presentadas, el término «evaluación» ha sido definido en múltiples ocasiones por otros tantos autores. Lázaro (1990:9-10) ofrece una reseña de veinte definiciones, que pone de manifiesto la pluralidad de acepciones con que cuenta el término en cuestión, de modo y manera que, en sí mismo, ya resulta confuso. Para Wolf (1990), que entiende la evaluación como una actividad técnica, propia de expertos, esta confusión es debida, en parte, a que los educadores suelen hacer uso de este término aún careciendo de conocimientos acerca de los instrumentos y procedimientos que se emplean en la evaluación educativa.

En un intento por clarificar lo que el término significa, Lázaro (1990:10), después de hacer un análisis de las veinte definiciones sobre evaluación aludidas anteriormente, destaca lo que cree que son las notas fundamentales de ésta, que están en línea con las señaladas por Orden Hoz (1991). En síntesis son las siguientes:

1. La evaluación abarca a todo hecho educativo (pudiendo entender por «hecho», a mi modo de ver, tanto objetos o entes abstractas como personas),
2. Supone tres fases claves, organizadas de forma secuencial:
 - 2.1. Recoger y sistematizar información,
 - 2.2. Valorar la información a todo hecho educativo, una vez comparada con un criterio o patrón,
 - 2.3. Tomar decisiones.

A esta última fase cabría añadir dos más; una derivada del concepto de evaluación de Beeby (1977), dirigida a la *planificación de acciones* que redunden en la mejora del sistema educativo, y otra referida a la *puesta en marcha y seguimiento de estas acciones* con la finalidad de *valorar las intervenciones* (Fernández Ballestrós, 1982; Pérez Juste, 1989).

Unido al concepto de evaluación, y no siempre bien diferenciado conceptualmente de éste, suele aparecer, como hemos visto, el término «valoración» (*assessment*), de modo que, en ocasiones frecuentes, cuando se habla de diagnóstico como evaluación, se observa un confusionismo terminológico y conceptual entre diagnóstico y este nuevo término que dificulta delimitar lo que es el diagnóstico. Por ello, vamos a analizar brevemente qué se entiende por «assessment o valoración».

1.2. Valoración (Assessment)

Siguiendo a Wolf (1990), el término «assessment» implica una valoración, comprobación o apreciación de características con fines diagnósticos, de clasificación o de gradación, que es realizada habitualmente por cada profesor con sus alumnos.

De acuerdo con Choppin (1990), el término *assessment* debe ser reservado para ser aplicado a personas, y está implícito en actividades relacionadas con exámenes, certificaciones, etc., que pueden ser realizadas con procedimientos formales o informales, y que se traducen en una calificación o gradación del sujeto para ubicarle o posicionarle dentro de una escala. Por ejemplo, comenta que puede emplearse este término cuando se analiza el rendimiento académico de un alumno, las actitudes de una persona para desempeñar un puesto de trabajo, la competencia profesional de un profesor, etc., cuyos resultados se traducen en la asignación de un número, un código o un comentario como, por ejemplo, «apto», «no apto», o «necesita mejorar»; estos procedimientos se asocian cada vez más, de modo incorrecto según Choppin, con la denominación «evaluación del alumno» o «evaluación del profesor».

Para este autor, la diferencia fundamental entre *assessment* y *evaluation* se centra en que mientras *assessment* hace referencia a una actividad rutinaria dentro del proceso de enseñanza-aprendizaje en la que están implicados la mayoría de los educadores, la evaluación es una actividad más técnica, ligada a la investigación y al desarrollo, y, por tanto, susceptible de ser realizada sólo por expertos, *con la finalidad última de mejorar la calidad del sistema educativo*; en este sentido, coincide con los planteamientos de Wolf (1990).

Los autores que siguen esta postura consideran que *el objeto de la evaluación* no son individuos o personas, sino objetos o entidades abstractas, tales como programas, currículos, métodos, recursos, organizaciones, etc., o grupos de individuos (poblaciones). Sin embargo, Popham (1988), haciendo un comentario sobre la confusión del empleo de estos términos, y el significado de *assessment*, llega a decir que es perfectamente lícito utilizar este último para denominar a los procedimientos que se ponen en marcha para valorar formalmente *programas* educativos de alcance nacional; en cambio, denomina «grading» a lo que Choppin (1990) entiende por *assessment*, centrado en individuos. En esta misma línea señalada por Popham se sitúa Santos Guerra (1990), quien entiende el *assessment* como una *comparación* de diversos *centros* a través de criterios homologados (1990:18):

«El *assessment* encierra una vertiente comparativa de la evaluación de los centros... Para ello habrá que tener en cuenta las condiciones de partida, los medios disponibles, las coordenadas en que se realiza una experiencia, etc. ... El *assessment* tiene la ventaja de que corrige la arbitrariedad y la falta de perspectiva que supone la valoración subjetiva de una experiencia ... En el *assessment* es importante la calidad original del dato que se va a someter al análisis y a la comparación. Y, desde luego, la validez de constructo de los instrumentos matemáticos que se utilizan para la medición y el análisis valorativo superior».

En cualquier caso, es claro que la función de *assessment* está siendo denominada cada vez con más frecuencia «evaluación del estudiante, del profesor, etc.».

1.3. Diagnóstico pedagógico

La delimitación del significado de Diagnóstico Pedagógico con respecto a los significados de los términos anteriores no resulta sencilla porque, de un modo u otro, éstos se encuentran implicados en la

realización de un diagnóstico. Esta implicación se advierte si repasamos los términos que aparecen en su definición, y los ponemos en relación con los conceptos de valoración (assessment) y evaluación (evaluation) que acabo de reseñar. En este sentido se observa que:

1. el Diagnóstico Pedagógico contempla actividades de assessment, relacionadas con la valoración de alumnos en un contexto de enseñanza-aprendizaje, ya que siguiendo a Álvarez Rojo (1984), Brueckner y Bond (1986), o en concreto a Gil Fernández (1991:151), se entiende al Diagnóstico Pedagógico como «un proceso que, mediante la aplicación de unas técnicas, específicas, permite llegar a un conocimiento más preciso del educando y orientar mejor las actividades de enseñanza-aprendizaje A través de la exploración de la actividad del escolar se intenta comprobar, describir y valorar los progresos del alumno en orden a la consecución de los objetivos educativos ...».
2. el Diagnóstico Pedagógico contempla también actividades de evaluación, tal como se desprende de la definición de Buisán y Marín (1984:13): «Incluye un conjunto de actividades de medición y **evaluación** de un sujeto (o grupo de sujetos) **o de una institución**». Asimismo, Lázaro (1986:81) define el Diagnóstico Pedagógico como «el conjunto de indagaciones sistemáticas utilizadas para conocer un hecho educativo con la intención de **proponer sugerencias y pautas perfectivas**. Desde esta inicial definición, cabe apreciar que el diagnóstico pedagógico debe superar el concepto de medida para, basándose tanto en ella como en valoraciones cualitativas, realizar una **evaluación** de todo el entramado escolar, bien se refiera a **producto, procesos, alumno o institución**». Pérez Juste (1990:17) añade, además, que se trata de una «**actuación técnica profesional, realizada con mentalidad científica, esto es con rigor**, encaminada a conocer a los educandos y a su ambiente...». Y el Diccionario de CC. de la Educación (1983, Vol.1:399), considera al diagnóstico como una «evaluación de una situación dada, juicio emitido sobre un hecho y su conjunto de circunstancias», que «presupone, además, una confrontación con la normatividad», es decir, con un criterio que en este caso se identifica con una norma aceptada; de manera que, una vez más, el concepto de Diagnóstico Pedagógico se acerca al *concepto de evaluación*.

Desde esta perspectiva, podríamos considerar que el Diagnóstico Pedagógico engloba tanto al «assessment» como a la «evaluation», y que tiene como *ámbitos de actuación* todos los que sean propios de la educación: sujetos, programas, instituciones docentes, familia, etc.).

1.4. Recapitulacion

La situación de confusión terminológica que venimos comentando ha sido denunciada por diversos estudiosos de nuestra disciplina (Pérez Juste, 1989; Lázaro, 1986, 1990), porque da lugar a una problemática no sólo de tipo conceptual, sino también práctico. Al dificultarse la definición de lo que es el Diagnóstico Pedagógico, se dificulta también la especificación de sus objetivos, la clarificación de sus funciones y la delimitación de sus ámbitos de actuación; así, por lo que respecta a estos dos términos, —diagnóstico y evaluación—, se pasa del empleo de uno a otro como si de sinónimos se tratara; Álvarez Rojo et al. (1984), por ejemplo, hablan de las «funciones evaluadoras/diagnósticas del profesor» (1984:38) propuestas por Benedito (1977), haciendo referencia a algunos aspectos que tradicionalmente han estado ligados al diagnóstico: «evaluación psicológica del alumno» (personalidad), «evaluación sociológica referida al comportamiento individual y social del alumno», y «evaluación del rendimiento escolar en las diferentes áreas del aprendizaje»; de otro lado, Fernández Ballesteros (1986:9) emplea también como sinónimos los términos «psicodiagnóstico» y «evaluación psicológica» o «evaluación escolar», y García Hoz y Pérez Juste (1984:64) comentan que «la evaluación tendrá que hacer referencia a la **capacidad** de los alumnos, que pudiéramos identificar con sus

aptitudes y, además, a un conjunto de otras variables, **personales** unas, ... **familiares** otras ... o **sociales** ...»; por su parte, Brueckner y Bond (1986:30) hablan, asimismo, de la «evaluación de la aptitud de un escolar».

Si nos ceñimos al uso que recomiendan algunos autores comentados para el término «evaluation», tendríamos que argumentar que éste ha sido empleado inapropiadamente en las situaciones descritas porque hacen referencia a personas y no a objetos físicos o a entidades abstractas.

De acuerdo con Pérez Juste (1989:36) este confusionismo entre «evaluación educativa» y «diagnóstico pedagógico», «probablemente se pueda explicar tanto por la evolución de ambos conceptos, que, progresivamente, han ido ampliando y enriqueciendo su significado, cuanto por las exigencias mismas del concepto de educación». Con estas palabras el autor se refiere al cambio que han experimentado ambos campos al pasar de un objetivo sumativo y curativo, respectivamente, a otro preventivo donde interesa potenciar al máximo el desarrollo personal.

2. HACIA UNA CLARIFICACIÓN CONCEPTUAL Y TERMINOLÓGICA

Un modo de clarificar si existen diferencias conceptuales relevantes entre Diagnóstico Pedagógico y Evaluación Educativa consiste en analizar las semejanzas y diferencias que se aprecian en las diversas definiciones emitidas para ambos. El apartado siguiente versa, precisamente, sobre las conclusiones obtenidas tras realizar este análisis.

2.1. Semejanzas y diferencias entre «diagnóstico pedagógico» y «evaluación educativa»

Comenzando con *las semejanzas*, que parecen ser las más abundantes, se observa en ambos, en primer lugar, una referencia a la recogida sistemática de información a través de técnicas que garanticen una cierta objetividad y precisión en la misma.

Por otra parte, tanto en las definiciones de diagnóstico como en las de evaluación, aparecen referencias a la interpretación y valoración de la información recogida, con el fin de facilitar una toma de decisiones. Esta toma de decisiones puede derivar en actuaciones diversas, y es aquí, precisamente, donde, a mi parecer, cabría establecer *algún tipo de distinción* entre Diagnóstico y Evaluación. De acuerdo con las definiciones analizadas de Diagnóstico Pedagógico, éste tiene como misión última «proponer sugerencias y pautas perfectivas» (Lázaro, 1986:81), ya sea sobre situaciones deficitarias, para su corrección y recuperación (Álvarez Rojo, 1984; Brueckner y Bond, 1986; Gil Fernández, 1991), o sobre situaciones no deficitarias, para su potenciación, desarrollo, o prevención de deficiencias (Pérez Juste, 1990). De esto se deduce que el diagnóstico pedagógico está enfocado siempre hacia la mejora, hacia el perfeccionamiento de su objeto de estudio —ya sea un producto o un proceso educativo—, y, por tanto, se contextualiza en un proceso perfectivo y de desarrollo —propio de la educación, en la que se inscribe—. En este sentido, comparte la finalidad de la *evaluación formativa* descrita por Scriven (1967), y se diferencia de la *evaluación sumativa* porque no tiene carácter sancionador.

Este carácter sancionador de la evaluación sumativa permite establecer, además, otra diferencia con el diagnóstico pedagógico. Si compartimos con Orden Hoz (1983) que este carácter sancionador condiciona el proceso y el producto educativo, haciendo que el objeto de estudio se adapte a los fines de la evaluación, entonces, hemos de distinguir al diagnóstico de esta evaluación; porque si, como dijimos antes, el diagnóstico tiene una finalidad perfectiva, de desarrollo, son sus fines, medios, y el propio proceso diagnóstico, los que han de adaptarse al objeto de estudio para llegar a conocerlo, y adecuar a él las pautas perfectivas (intervenciones), y no a la inversa. Quizás, por esta razón, Ander

Egg (1987) considere que *para hacer un diagnóstico se necesiten ciertos conocimientos del campo de la planificación.*

Una perspectiva más a comentar que puede resultar clarificadora de las diferencias analizadas, pero ahora desde otro punto de vista, la encontramos en Buisán y Marín (1987:14), quienes establecen una distinción entre ambos conceptos basada en la *cantidad de técnicas que es necesario emplear para poder realizar un diagnóstico pedagógico o una evaluación*: «el diagnóstico implica una labor de *síntesis e interpretación* de las informaciones recogidas mediante técnicas diversas. Podremos hacer evaluación con una sola prueba, pero nunca podremos hacer un diagnóstico con una sola prueba o información».

Las diferencias entre ambos conceptos, quizás deban buscarse también *en relación al objeto* analizado. En este sentido, hemos visto que diversos autores han restringido el concepto de «*evaluation*» a objetos físicos o entes abstractas, excluyendo el elemento humano, que ha sido reservado al concepto de «*assessment*», —valoración—, y que ha sido, tradicionalmente, objeto de estudio directo del Diagnóstico Pedagógico. Pero, en tanto en cuanto el diagnóstico tenga el apellido «*pedagógico*», es mi postura considerar que deberá incluir como objeto de estudio no sólo al individuo que aprende o se forma como tal, sino también al contexto en el que este proceso tiene lugar (Fernández Ballesteros, 1982), ya sea la institución escolar con su organización, métodos didácticos, personal y recursos determinados (Álvarez Rojo, 1984; Brueckner y Bond, 1986; Buisán y Marín, 1987; Pérez Juste, 1989), o la comunidad educativa que rodea a dicha institución (familia, barrio, etc.). En este sentido, y teniendo en cuenta el objeto, cabe decir, como ya hemos visto también antes, y sin ánimo de delimitar conclusiones definitivas, que *el diagnóstico es un concepto más amplio que el de «evaluation».*

Este punto nos conduce a un debate ya iniciado referido al *proceso diagnóstico*, en el que se intenta dar respuesta a la pregunta siguiente.

2.2. ¿Es el diagnóstico una fase del proceso de evaluación, o, por el contrario, envuelve a ésta en un proceso más amplio?

Esta pregunta conduce a otras como las siguientes: ¿cabe pensar que evaluación y diagnóstico pedagógico son procesos independientes?; en el caso de que la respuesta fuera negativa, ¿cabe hablar de procesos interrelacionados, o más bien de todos y partes?; en este último caso ¿qué proceso envuelve a qué otro?, ¿es el diagnóstico pedagógico tan sólo una fase del proceso de evaluación, o, por el contrario, es ésta la que forma parte del proceso diagnóstico?

Parece complejo dilucidar este punto; por ello, no sorprende encontrar entre la bibliografía consultada posturas variadas, defendiendo una u otra opción. Por ejemplo, mientras Pawlik (1980) o Lázaro (1986, 1990) sostienen que el proceso diagnóstico no concluye en la valoración emitida como fruto de la evaluación, sino que «se podría considerar que el diagnóstico incluye el acto evaluador, ya que supone un proceso permanente de control de la intervención, pasando desde las referencias situacionales hasta la toma de decisiones de cualquier variable o entidad de tipo educativo» (Lázaro, 1990:11), Pérez Juste (1989) considera que el diagnóstico —que tiene para este autor una mera función de conocimiento y descripción de limitaciones y posibilidades, de aspectos positivos y negativos de la realidad pedagógica—, es la fase inicial del proceso evaluador: «nuestra concepción supone la inclusión del diagnóstico como una dimensión y etapa integrada en el concepto de evaluación, la etapa inicial o evaluación inicial» (1989:38). «La evaluación precisará del diagnóstico para la realización de los pronósticos que permitan una actuación preventiva y que faciliten los juicios de valor de referencia personalizada ..., además, claro está, de para personalizar el proceso educativo...» (1989:37). No obstante, en el texto compartido por este autor con García Hoz (1984) admite la continuidad de la

función diagnóstica a lo largo de todo el proceso evaluador, incluyéndola incluso en la evaluación final.

Un aspecto a destacar en relación a los comentarios efectuados por éstos y otros autores es la referencia al *pronóstico* que ha de llevar asociado todo diagnóstico (Ander Egg, 1987; Pérez Juste, 1989). A este respecto, Gil Fernández (1991:151) comenta: «el diagnóstico no es completo sin un pronóstico y un plan de actuación formulados a la vista del análisis de los datos obtenidos. Este pronóstico es el punto de partida y de llegada del proceso de rehabilitación derivado del diagnóstico».

Por su parte, Álvarez Rojo y colaboradores (1984:23), refiriéndose a uno de los aspectos que garantiza la eficacia de un programa de enseñanza correctiva, sostienen que «es necesario establecer un sistema de evaluación que permita determinar los progresos del alumno desde la situación inicial revelada por el diagnóstico hasta la aplicación concreta de cada una de las partes del programa», dando a entender que el diagnóstico es, como apuntaba Pérez Juste (1989), la fase inicial del proceso evaluador y que, por tanto, ésta envuelve a aquél.

Dada la diversidad de posturas, cada una con su lógica, no resulta fácil optar por una u otra; desde mi punto de vista, diagnóstico y evaluación forman parte, indefectiblemente, de *un mismo proceso*, porque no puede haber diagnóstico sin valoración, es decir, sin evaluación. El proceso de diagnóstico desemboca en una interpretación de la realidad estudiada, pero esta interpretación ha de estar basada en una valoración previa de la misma, que surge, a su vez, del grado de ajuste existente entre las características de dicha realidad y las de un criterio o patrón que se toma como punto de comparación. En este sentido, si hubiera que destacar alguna diferencia entre diagnóstico y evaluación en relación al proceso, señalaría que ésta se circunscribe únicamente al momento en que tiene lugar una u otra actividad dentro del mismo proceso: la evaluación o valoración es anterior a la interpretación diagnóstica.

3. CONSIDERACIÓN FINAL

Tras efectuar estas matizaciones entre diagnóstico y evaluación no resulta extraño observar que en muchos casos se utilicen ambos términos como intercambiables, dada la necesaria conexión que existe entre ambos; es posible que no resulten fácilmente diferenciables si no se realiza una reflexión sobre ellos que permita distinguir los matices que hacen diferente a un concepto de otro.

Lo que con claridad les diferencia es el número de definiciones que se encuentran en los textos para uno y otro término; mientras numerosos autores dentro y fuera de nuestro país se han ocupado de analizar y definir el concepto de «evaluación educativa», encontramos muy pocas referencias bibliográficas donde se defina lo que es el «diagnóstico pedagógico», quizás porque tradicionalmente se le ha entendido ligado al psicodiagnóstico y al campo médico, y sólo en los últimos años se le está considerando en nuestro país como disciplina con entidad propia dentro de las CC. de la Educación.

Un ejemplo clarificador del escaso empleo que se hace del término «diagnóstico pedagógico» frente al de «evaluación» se obtiene si se repasa el título de las ponencias y comunicaciones presentadas en el V Seminario organizado por la asociación AIDIPE, celebrado en Murcia (1990), titulado «Metodologías en el diagnóstico y evaluación en los procesos de intervención educativa». Ya en el título se advierte una intención de diferenciar entre diagnóstico y evaluación, dado que aparecen unidos por la conjunción «y», pero lo que resulta llamativo es que en ninguna de estas presentaciones se ofrece una definición del término diagnóstico, y prácticamente, éste no se incluye en los títulos de las mismas, frente al término «evaluación», que aparece en la mayoría de ellas.

4. REFERENCIAS BIBLIOGRÁFICAS

- ÁLVAREZ ROJO, V. et al. (1984): *Diagnóstico Pedagógico*. Sevilla, Alfar.
- ANDER EGG, E. (1987): *Investigación y diagnóstico para el trabajo social*. Buenos Aires, Humanitas.
- BEEBY, C.E. (1977): «The meaning of evaluation», *Current issues in evaluation*, 4. Wellington, New Zealand, Department of Education.
- BRUECKNER, L. J. & BOND, G. L. (1986): *Diagnóstico y tratamiento de las dificultades en el aprendizaje*. Madrid, Rialp.
- BUISAN, C. y MARÍN, M. A. (1984): *Bases diagnósticas y prácticas para el diagnóstico pedagógico*. Barcelona, PPU.
- BUISAN, C. y MARÍN, M. A. (1987): *Cómo realizar un diagnóstico pedagógico*. Barcelona, Oikos-Tau.
- CHOPPIN, B. H. (1990): «Evaluation, assessment, and measurement», en H. J. Walberg & G. D. Haertel (Eds.): *The International Encyclopedia of Educational Evaluation*. Oxford, Pergamon Press.
- FERNÁNDEZ BALLESTEROS, R. et al. (1982): *Evaluación de contextos*. Murcia, Servicio de Publicaciones de la Universidad de Murcia.
- GIL FERNÁNDEZ, P. (1991b): «Diagnóstico Pedagógico», en Santillana (Ed.): *Léxicos de Ciencias de la Educación. Tecnología de la educación*. Madrid, Santillana.
- FERNÁNDEZ BALLESTEROS, R. (1986): *Psicodiagnóstico*, Vol. 1. Madrid, UNED.
- GARCÍA HOZ, V. & PÉREZ JUSTE, R. (1984): *La investigación del profesor en el aula*. Madrid, Escuela Española.
- LÁZARO MARTÍNEZ, A. J. (1986): *Orientación y educación especial*. Madrid, Anaya.
- LÁZARO MARTÍNEZ, A. J. (1990): «Problemas y polémicas en torno al diagnóstico pedagógico», *Bordón*, 42(1):7-15.
- NEVO, D. (1990): «Role of the evaluator», en H. J. Walberg y G. D. Haertel (Eds.): *The International Encyclopedia of Educational Evaluation*. Oxford, Pergamon Press.
- ORDEN HOZ, A. (1983): «La investigación sobre evaluación educativa», *Revista de Investigación Educativa*, 2:240-258.
- ORDEN HOZ (1989): Prólogo al libro de J. M. García Ramos. *Bases pedagógicas de la evaluación*. Madrid, Síntesis.
- ORDEN HOZ, A. (1991a): «La evaluación de programas en el contexto educativo», Ponencia presentada al Seminario sobre Evaluación de Programas Educativos. Madrid.
- PAWLIK, K. et al. (1980): *Diagnosis del diagnóstico*. Barcelona, Herder.
- PÉREZ JUSTE, R. & GARCÍA RAMOS, J. M. (1989): *Diagnóstico, evaluación y toma de decisiones*. Madrid, Rialp.
- PÉREZ JUSTE, R. (1990): «Recogida de información en el diagnóstico pedagógico», *Bordón*, 42(1):17-29.
- POPHAM, W.J. (1988): *Educational Evaluation*. New Jersey, Prentice Hall.
- SANTOS GUERRA, M. A. (1990): *Hacer visible lo cotidiano*. Madrid, Akal.
- WALBERG, H. J. & HAERTEL, G. D. (Eds.) (1990): *The International Encyclopedia of Educational Evaluation*. Oxford, Pergamon Press.
- WOLF, R. M. (1990): *Evaluation in Education. Foundations of competency assessment and program review*. New York, Praeger.

ESTUDIO LONGITUDINAL DEL TEMPERAMENTO INFANTIL DE 0 A 12 MESES: RESULTADOS PRELIMINARES DE UNA APROXIMACIÓN METODOLÓGICA INTERDISCIPLINAR*

Francisco A. García-Sánchez,
Facultad de Educación. Universidad de Murcia
Julio Pérez-López y José A. Carranza
Facultad de Psicología. Universidad de Murcia

INTRODUCCIÓN

Este trabajo resume algunos resultados preliminares de un estudio longitudinal acerca de la detección y estabilidad de los rasgos temperamentales en el niño, desde el nacimiento hasta los dos años de edad. El concepto de temperamento hace referencia a las pautas de organización fisiológica del niño (probablemente determinadas genéticamente) que permiten predecir en él una forma única y exclusiva de pensar, sentir y actuar de manera más o menos constante (Rothbart y Derryberry, 1981). Se trata del estilo, quizás innato, con el que el niño se enfrenta a las experiencias y relaciones con el ambiente y que esas mismas experiencias y relaciones ambientales irán matizando. El temperamento está mediando en la forma con que el niño interpreta el mundo exterior (Thomas y Chess en Goldsmith et al. 1987). Por ello, el temperamento está en la base (junto con el ambiente) de la formación de la personalidad y de los futuros estilos cognitivos o estrategias de procesamiento de la información; elementos ambos que tradicionalmente se consideran claves para la generación en el niño de un estilo propio de aprendizaje en modelos como los de Biggs, Curry o Entwistle (ver revisión de Hernández-Pina, Iglesias y Serrano, 1990).

En el estudio del temperamento infantil podemos distinguir básicamente dos aproximaciones metodológicas: el uso de cuestionarios cumplimentados por los padres y la observación directa de la conducta del niño en casa y/o en el laboratorio. Normalmente estas dos aproximaciones se han utilizado separadamente (p.ej., Thomas y Chess, 1977; Matheny y Wilson, 1981) y, en ocasiones, se ha intentado correlacionar una de ellas con medidas psicofisiológicas que reflejen la activación del sistema nervioso (p.ej., Stifter y Fox, 1990). Sin embargo, se ha hecho hincapié en la necesidad de

* Esta investigación ha sido parcialmente financiada por la Dirección General de Educación y Universidad de la Consejería de Cultura, Educación y Turismo de la Comunidad Autónoma de la Región de Murcia (PSH 89/51) y por la Dirección General de Investigación Científica y Técnica (DGICYT) (PB90-0309).

combinar estas diferentes metodologías para facilitar la obtención de una definición y conceptualización operativa del temperamento (ver revisión de Bates, 1989), y hacerlo desde una perspectiva longitudinal para poder analizar la posible estabilidad normativa de las diferencias individuales (Buss y Plomin, 1984; Rothbart, 1989).

La unificación de las tres estrategias metodológicas en el estudio del temperamento infantil es, precisamente, el objetivo de nuestra investigación interdisciplinar. Pretendemos evaluar, desde una perspectiva longitudinal, la concordancia entre los parámetros del temperamento infantil obtenidos a partir de cuestionarios, de la observación directa de la conducta del niño en laboratorio y del registro de su actividad cardíaca en las mismas situaciones de laboratorio.

MÉTODO

Sujetos

La muestra final estuvo formada por 60 recién nacidos a término (29 niñas y 31 niños), sin complicaciones pre ni postnatales, cuyas madres se ofrecieron voluntariamente para participar en el estudio (todas acudían a clases de preparación al parto en el Hospital Virgen de la Arrixaca de Murcia). La edad media de las madres fue de 27.5 años ($dt= 4.64$; rango= 21-42 años), con nivel socioeconómico medio.

Instrumentos y aparatos

Para obtener la valoración que los padres hacen del temperamento de sus hijos se emplearon traducciones del Infant Behavior Questionnaire (IBQ) de Rothbart (1981), aplicable hasta los 9 meses de edad, y del Toddler Behavior Assessment Questionnaire (TBAQ) de Goldsmith (1987), aplicable a partir de los 12 meses. En ellos se pregunta acerca de conductas de los hijos en situaciones cotidianas de juego, comida, baño, sueño, etc. Reflejan seis dimensiones temperamentales: nivel de actividad, risa/sonrisa, miedo, angustia ante las limitaciones, facilidad para tranquilizarse (sólo hasta los 9 meses) y duración de la orientación.

Para obtener una medida conductual directa del temperamento infantil se emplearon las Tareas Evolutivas y Escalas de Valoración para la Medida del Temperamento Infantil en el Laboratorio de Matheny y Wilson (1981). Consisten en una serie de tareas o viñetas (con orden de presentación, descansos e interrupciones preestablecidos), que suponen una secuencia estructurada de interacciones, vinculadas con la edad del niño, que se presentan como retos o desafíos a los que el niño ha de enfrentarse. Las interacciones se establecen, según las tareas, con el experimentador, con objetos físicos o con situaciones frustrantes. Las sesiones se grababan en vídeo y posteriormente eran analizadas por observadores entrenados independientes, quienes puntuaron las siguientes dimensiones temperamentales: tono emocional, actividad, orientación social hacia el experimentador, atención y vocalizaciones. El índice de fiabilidad entre observadores fue de 0.91.

Para obtener una medida periférica de la activación del sistema nervioso se registró la actividad cardíaca a través de un Electrocardiógrafo BEXEN monocanal (con velocidad de registro de 25 mm/seg) por medio de tres electrodos pediátricos colocados en el pecho del niño. Se cuantificó manualmente el período cardíaco como intervalo de tiempo (medido en milisegundos) transcurrido entre dos ondas cardíacas. A partir de esta cuantificación se obtuvieron valores de frecuencia cardíaca (FC) (en latidos por minuto) y de variabilidad del período cardíaco (desviación típica y rango, en milisegundos).

Procedimiento

Todas las sesiones de evaluación se llevaron a cabo, con periodicidad trimestral (3, 6, 9 y 12 meses de edad), en la Sala de Observación y Registro de la Escuela Infantil de Guadalupe de la Comunidad Autónoma de la Región de Murcia. El procedimiento de evaluación fue idéntico en todos los casos, salvo las tareas del laboratorio que, por requerimientos evolutivos, variaron a partir de los 9 meses de edad. Su duración media fue de 70 minutos, comenzando con un periodo de familiarización a la situación de laboratorio (10 min.), en el que se producía la primera toma de contacto entre el examinador y el niño. Tras él, se colocaban los electrodos y se procedía al registro de la actividad cardíaca (3 min.), con el niño en posición de decúbito supino (periodo de línea base). A continuación comenzaba la administración de las tareas evolutivas en el orden propuesto por Matheny y Wilson (1981). Mientras no fuese necesaria su presencia, los padres se retiraban a la sala de control donde cumplimentaban los cuestionarios de temperamento infantil.

RESULTADOS Y CONCLUSIONES

La estabilidad de cada una de las medidas del temperamento fue estudiada calculando matrices de correlación de Pearson entre las distintas edades, cuyos resultados se resumen en la Tabla 1. En prácticamente todas las dimensiones del temperamento obtenidas en laboratorio se observa una gran estabilidad entre los 9 y 12 meses. La falta de una clara estabilidad antes de estas edades puede deberse

TABLE 1
 COEFICIENTE DE CORRELACIÓN DE PEARSON ENTRE EDADES PARA LA ACTIVIDAD
 CARDÍACA Y LAS DIMENSIONES DE TEMPERAMENTO OBTENIDAS EN LABORATORIO Y A
 TRAVÉS DE CUESTIONARIOS

	COMPARACIONES ENTRE MESES			
	1 vs 3	3 vs 6	6 vs 9	9 vs 12
DIMENSIONES EN LABORATORIO				
Tono emocional	—	.45***	.24	.52***
Actividad	—	-.08	.22	.14
Atención	—	.47***	.20	.46***
Orientación Social	—	.00	.13	.47***
Vocalizaciones	—	.11	.37**	.45***
DIMENSIONES EN CUESTIONARIOS				
Actividad	.71***	.61***	.52***	.47***
Ansiedad ante limitaciones	.36*	.59***	.62***	.29*
Miedo	.61***	.37**	.57***	.38**
Duración Orientación	—	.55***	.61***	.31*
Risa/sonrisa	.47***	.47***	.52***	.63***
Facilidad Tranquilizarse	.41**	.15	.63***	—
ACTIVIDAD CARDÍACA				
Desviación típica del período cardíaco	—	.20	.33*	.21
Rango del período cardíaco	—	.17	.31*	.18
Frecuencia cardíaca	—	.24	.04	.39**

a desarrollos endógenos de los niños que ocurren en torno a los 7/8 meses de edad (Bates, 1989). Por su parte, las dimensiones de temperamento obtenidas a partir de cuestionarios mostraron correlaciones significativas a todas las edades, apoyando la estabilidad que defiende Rothbart (1989) en las percepciones maternas del temperamento de sus hijos.

Las medidas de variabilidad del período cardíaco (desviación típica y rango) sólo mostraron estabilidad entre los 6 y 9 meses (ver Tabla 1), mientras que la FC sólo fue estable entre los 9 y 12 meses. Probablemente, a edades más bajas, la propia inmadurez fisiológica del niño impide una mayor estabilidad. En este sentido, un MANOVA realizado para comparar los valores de FC en las distintas edades resultó significativo ($F_{(3,132)}=26.32, p<.001$). El análisis de tendencias indicó el esperado descenso de la FC con el aumento de la edad de los niños (Stifter y Fox, 1990). No obstante, este mismo análisis de tendencias señaló un aumento de la FC a los 12 meses. El futuro análisis de los distintos registros electrocardiográficos obtenidos a lo largo de la sesión de evaluación determinará si, como creemos, este resultado se debe a un mayor estado de ansiedad en el niño ante la situación de laboratorio.

Para comprobar hasta que punto existen relaciones (o se pueden establecer predicciones) entre las observaciones directas de los niños, las medidas indirectas y las psicofisiológicas, se efectuaron análisis de correlación canónica entre las distintas medidas. Resumiendo sus resultados podemos decir que a los 12 meses, la percepción que del temperamento de sus hijos tienen las madres explica el 36.8% de la varianza del temperamento medido en las situaciones de laboratorio ($X^2_{(30)}=59.76; p<.01$). A partir de esa edad las madres perciben adecuadamente el temperamento de sus hijos, por lo que, como señala Mebert (1991), estaría justificado el empleo de cuestionarios del temperamento infantil. Por su parte, a la edad de 3 meses las diferencias en la actividad cardíaca explican el 16.86% de la varianza del temperamento medido en laboratorio ($X^2_{(30)}=25.37, p=.04$), mientras que esta asociación desaparece en las demás edades. Este último resultado, junto con los anteriormente descritos, indican que el temperamento a los 3 meses puede estar influenciado por rasgos psicofisiológicos internos, mientras que a partir de los 12 meses son las variables contextuales y de socialización las que predominan sobre lo biológico.

Así pues, las madres mantienen estable a lo largo de nuestro estudio la percepción que tienen del temperamento de sus hijos. Sin embargo, esta percepción sólo converge con las dimensiones del temperamento obtenidas en laboratorio a los 12 meses, mientras que las variables fisiológicas pierden esa convergencia a partir de los 6 meses. De todo ello parece deducirse que la manifestación del temperamento del niño que observamos en laboratorio va adaptándose a las expectativas que del mismo tienen sus padres, alejándose incluso de su posible correlato fisiológico. En un futuro próximo, nuestros propios datos pueden confirmar esta conclusión si, efectivamente, a partir de los 12 meses seguimos encontrando esa asociación entre las percepciones maternas y las medidas del temperamento infantil en laboratorio.

BIBLIOGRAFÍA

- BATES, J. E. (1989): Concepts and measures of temperament. En G. A. Kohnstamm, J. E. Bates y M. K. Rothbart (Eds.): *Temperament in childhood* (pp. 3-27). Nueva York: Wiley.
- BUSS, A. H. y PLOMIN, R. (1984): *Temperament: Early developing personality traits*. Hillsdale, NJ: Erlbaum.
- GOLDSMITH, H. H. (1987): *Toddler behavior assesment questionnaire*. Universidad de Oregón.
- GOLDSMITH, H. H., PLOMIN, R., ROTHBART, M. K., THOMAS, A., CHESS, S., HINDE, R. A. y MCCALL, R. B. (1987): Rountable: What is temperament? Four approaches. *Child Development*, 58, 505-529.

- HERNÁNDEZ PINA, F., IGLESIAS, E. y SERRANO, F. J. (1990): Enfoques de aprendizaje universitario como base para el diagnóstico de necesidades. *Revista de Investigación Educativa*, 16, 239-253.
- MATHENY, A. P. y WILSON, R. S. (1981): Developmental tasks and rating scales for the laboratory assessment of infant temperament. *JSAS Catalog of Selected Documents in Psychology*, 11, 81 (Manuscrito núm. 2.367).
- MEBERT, C. (1991): Dimensions of subjectivity in parents' ratings of infant temperament. *Child Development*, 62, 352-361.
- ROTHBART, M. K. (1981): Measurement of temperament in infancy. *Child Development*, 52, 569-578.
- ROTHBART, M. K. (1989): Temperament and development. En G. A. Kohnstamm, J. E. Bates y M. K. Rothbart (Eds.): *Temperament in childhood* (pp. 187-247). Nueva York: Wiley.
- ROTHBART, M. K. and DERRYBERRY, D. (1981): The development of individual differences in temperament. In M. LAMB and A. L. BROWN (Eds.): *Advances in developmental psychology* (vol. 1, pp. 37-86). Hillsdale, NJ: Erlbaum.
- STIFTER, C. A. y FOX, N. A. (1990): Infant reactivity: Physiological correlates of newborn and 5-month temperament. *Developmental Psychology*, 26, 582-588.
- THOMAS, A. y CHESS, S. (1977): *Temperament and development*. Nueva York: Brunner/Mazel.

UNA EXPERIENCIA DE DIAGNÓSTICO DE LOS PROCESOS DE ENSEÑANZA-APRENDIZAJE EN LA UNIVERSIDAD A TRAVÉS DE LA INVESTIGACIÓN-ACCIÓN

Soledad Romero Rodríguez (coord.), Dolores Ballester Angulo, Dolores Flores García, Juan García Murillo, Luis Giráldez Sánchez, M^a Dolores Macipe Fernández, Encarnación Rubio Rodríguez y Amparo Sánchez Sánchez
Universidad de Sevilla

1. INTRODUCCIÓN

Los procesos de Enseñanza-Aprendizaje siguen siendo en nuestras universidades, en líneas generales, demasiado tradicionales, y acusan de falta de flexibilidad e innovaciones. Consideramos que una metodología de investigación-acción participativa donde cada miembro del grupo-clase se implique en los procesos de diagnosis, planificación, acción y reflexión, favorecería, por propia definición, la co-gestión de los procesos educativos. Por otra parte, se respondería a la necesidad de formar profesionales no sólo con «cabezas llenas» y, si acaso —en muchas ocasiones ni siquiera esto se da— una maleta llena de técnicas y procedimientos, sino también con actitudes —de escucha, trabajo en equipo, reflexión y autoanálisis— sin las cuales no podríamos hablar de un verdadero desarrollo profesional. Nuestra pretensión en esta comunicación es presentar a grandes rasgos las características de los procesos que se han dado en nuestra —de la profesora y los alumnos— asignatura de Diagnóstico Pedagógico durante el curso 1992-93. Presentamos un análisis provisional pues aún no hemos dado por terminado este proceso.

2. MARCO TEÓRICO Y EXPERIENCIAL

El actual sistema educativo universitario ha de pasar necesariamente por un proceso de democratización y gestión conjunta de administradores, profesores y alumnos. Consideramos que para que esto funcione, tanto profesores como alumnos han de flexibilizar sus papeles tradicionales.

Nuestra experiencia ha pretendido ser un intento de poner en práctica estos planteamientos en la asignatura de diagnóstico Pedagógico de la licenciatura en Pedagogía. El grupo, constituido por siete alumnos y la profesora decide emprender un proyecto de diagnóstico sobre su propio proceso de enseñanza-aprendizaje. A través de este proyecto irá surgiendo un modelo didáctico de co-gestión, fruto de la metodología usada para el diagnóstico: la Investigación-Acción.

Esta experiencia ha permitido que el grupo adquiera una mayor conciencia de su propia realidad personal y profesional (como alumnos y profesionales en prácticas). En este sentido, nos acercamos al

aspecto educativo que Gauthier y Baribeau (1981, en Goyette y Lessard-Herbert, 1988) consideran central la investigación-acción al optar por «una investigación-acción que pone un acento prioritario en el proceso de aprendizaje del conjunto de personas que están implicadas en ella» (Gauthier y Baribeau, en Goyette y Lessard-Herbert, 1988:107).

Gracias a la metodología de Investigación-Acción utilizada en el desarrollo del proceso de Enseñanza-Aprendizaje, éste permite hablar de un nuevo perfil de alumno: alumno creador y activo, constructor del conocimiento, solidario y comprometido con el grupo al que pertenece y profundamente implicado en el desarrollo del proceso de enseñanza y aprendizaje, ya que lo llega a identificar como «suyo». Igualmente se favorece un perfil de profesor formador, animador, investigador y dinamizador, acompañante del alumno. Por lo tanto se hace del grupo la base del proceso de enseñanza-aprendizaje (López Górriz, 1991).

3. PROCESO DIAGNÓSTICO

Antes de empezar queremos dejar constancia de lo difícil que resulta cualquier intento de sistematizar el proceso que hemos llevado a cabo. En primer lugar, hemos de considerar que existen dos momentos diferentes que se dan simultáneamente —*acción* y *reflexión*— a lo largo de cada una de las fases del proceso (Ander-Egg (1990) habla de procesos de acción-reflexión). Esta simultaneidad de los dos momentos nos impide organizar linealmente el proceso seguido; no obstante, intentaremos exponerlo de la manera más clara posible.

Nuestro proceso de diagnóstico se inicia a raíz de una propuesta por parte de la profesora de la asignatura, enunciada de la siguiente forma: realizar un diagnóstico sobre el proceso de enseñanza-aprendizaje que se estaba llevando a cabo en nuestra clase a través de la metodología de investigación-acción.

En un principio, el grupo (sus miembros alumnos) mostró una actitud apática debida, en parte, a la falta de costumbre como participante en procesos de co-gestión. Sólo más tarde, gracias a una serie de factores, tanto internos como externos, se da pie a que este proceso comience a integrarse en la vida del grupo. Entre los factores que facilitaron esta nueva actitud podemos destacar: a) Factores internos: la clase entra en una dinámica que posibilita a los alumnos el ser participantes activos y no meros receptores pasivos de un proceso de enseñanza-aprendizaje; b) Factores externos: nos encontramos en ese momento más relajados puesto que ya se han celebrado los primeros exámenes; por otra parte, recibimos una fuerte influencia de otra profesora especialista en esta metodología de investigación.

«Hoy **** nos ha hablado de la crisis que conlleva implicarse en investigación-acción, su apasionamiento me ha hecho pensar, me ha provocado a actuar, hay cosas con las que no estoy de acuerdo, gente, métodos prehistóricos; en ese sentido, quizás sea interesante lo que nos proponen en esta asignatura y realizar un diagnóstico sobre el proceso de enseñanza-aprendizaje a través de la investigación-acción».

Tras establecer unos objetivos generales, que han sido continuamente revisados, llegamos a establecer un contrato, un acuerdo que dejara claros los límites de la investigación, el grado de implicación de los miembros, la temporalización de la investigación, las técnicas de recogida y análisis de datos, etc. Los objetivos inicialmente planteados (finalidad del diagnóstico-intervención) se concretaron en: 1) análisis del proceso de enseñanza aprendizaje que se daba en el aula; b) facilitar el desarrollo profesional-personal de todos los participantes (profesora y alumnos). Como señalamos anteriormente, estos objetivos fueron revisados de forma continua a lo largo de todo el proceso. Estas revisiones implicaron una constante profundización en la definición consensuada de los objetivos.

Como instrumentos de recogida de información optamos por utilizar fundamentalmente el diario. Cada miembro del grupo llevaba su propio diario con la finalidad de poder obtener información válida en el contraste de los mismos. Asimismo, desarrollamos sesiones de grupo cuyo objetivo era renegociar, consensuar, reflexionar acerca del trabajo que estábamos realizando, de una parte, y, por otra, realizar un análisis reflexivo en torno a los diferentes elementos del proceso de enseñanza-aprendizaje (tipo de actividades, desarrollo de las clases, rol de la profesora, implicación de los alumnos...). Estas sesiones fueron grabadas en audio.

Para el análisis de la información recogida hemos seguido el siguiente procedimiento: (a) Determinación de las dimensiones a analizar, a través de un proceso inductivo y consensuado. Las dimensiones derivadas de este proceso fueron las siguientes: Organización (asistencia a clase, espacios, tiempo), Planificación (metodologías de registro, objetivos), metodología didáctica, preparación de las clases, desarrollo de las clases, aprendizajes, roles, grupo e implicación en la investigación (acuerdos), todo ello con el fin de alcanzar los objetivos fijados anteriormente; (b) Revisión individual de los diarios y las transcripciones de las sesiones en función de las dimensiones, aunque de forma abierta a otras nuevas que pudiesen aparecer; (c) Consenso del análisis realizado e interpretación de los resultados en sesiones de grupo.

4. PRIMERAS IMPRESIONES DERIVADAS DEL ANÁLISIS

En el momento de redacción de esta comunicación nos encontramos en la fase de interpretación consensuada de las informaciones recogidas. Es por ello por lo que no podemos presentar más que una descripción y un primer intento de interpretación de la experiencia.

En un principio la asistencia a clase era media-baja. A medida que fue transcurriendo el curso y fuimos tomando conciencia del proceso en el que íbamos entrando, la asistencia aumentó considerablemente, hasta completarse en la última parte del curso, lo que interpretamos, en función de nuestros debates y reflexiones como un indicador de la progresiva implicación de los miembros del grupo. Es entonces cuando se detecta, cuando vivimos la conciencia de grupo, que se ve favorecida por la organización espacial de las clases. Hemos de señalar en este sentido, que el espacio de trabajo lo consideramos como un factor condicionante de este tipo de procesos. En un principio las clases se desarrollaban en una gran aula, con filas de mesas y una mesa para el profesor colocada sobre una tarima (aula tradicional). Posteriormente, decidimos buscar un aula-seminario que favoreciera nuestra interacción.

«Me siento un poco incómoda en este aula, ya que es demasiado grande y nos hemos tenido que sentar en filas, yo no sé dónde ponerme para facilitar la interacción» (Diario de la profesora).

«Hoy nos hemos ido al despacho. El aula XVIII no sirve para el tipo de trabajo que llevamos, ¿por qué serán todas las aulas de este tipo en esta Facultad, precisamente en esta Facultad» (Diario de un alumno).

Al ir tomando conciencia de grupo, los roles sufren una transformación, destacando el rol de la profesora que pasa de ser directiva a ser animadora, orientadora, facilitadora.

«No se trata de una clase convencional en la que el profesor lleva el peso de la misma, sino que los alumnos participan, dan su opinión, puntos de vista y juicios de valor, es decir, se trata de una clase democrática y abierta (...) Está ***, que, aunque a veces tenga que llevar el peso de la clase, casi siempre se convierte en una compañera más, se pone a nivel nuestro» (Diario de un alumno).

El grupo, con respecto a sus roles, trata de convertirse en más autónomo en el proceso de enseñanza-aprendizaje, y la máxima autonomía llega a través de la reflexión grupal de nuestros diarios. La implicación se observa no sólo en la asistencia, sino también en el compromiso hacia la metodología que estamos llevando y a la no-directividad de la profesora.

El aprendizaje de la asignatura viene a través del compartir las experiencias que tienen los alumnos en sus trabajos de campo fuera del aula (realización de un proceso diagnóstico en distintos centros educativos). Con esto se modifica la metodología tradicional, a la que estamos tan acostumbrados. Por otra parte, se construye el conocimiento en grupo, con una riqueza que no se hubiese alcanzado si el trabajo hubiese sido individual.

«Comienzo yo y cuento mi proceso de observación, trato de mirar a todos cuando hablo. Nos hemos comprometido a intercambiar». (Diario de un alumno).

El aprendizaje de la asignatura no ha sido sólo de carácter teórico-conceptual, sino también de carácter procedimental (a través de las prácticas diagnósticas fuera y dentro del aula) y actitudinal. Algunas de las actitudes que se han desarrollado son: sentido crítico, escucha, expresión en público, tolerancia. Por otra parte, a nivel profesional se han conseguido aprendizajes relativos a la creencia en la posibilidad de la innovación educativa y se ha dado un giro en la concepción del trabajo del pedagogo que tenían los alumnos.

«He intervenido más que de costumbre, y esto para mí supone romper con la timidez que me caracteriza» (Diario de un alumno)

Han sido muchas las limitaciones (temporales, institucionales, de resistencia al cambio) que hemos tenido que afrontar. Aún así, hemos conseguido llegar a un proceso de co-gestión.

BIBLIOGRAFÍA

- ANDER-EGG, E. (1990): *Repensando la investigación-acción participativa*. Vitoria. Departamento de trabajo y Seguridad Social.
- GOETZ y LECOMPTE (1988): *Etnografía y diseño cualitativo en investigación educativa*. Madrid. Morata.
- GOYETTE, G. y LESSARD-HERBERT, M. (1988): *La investigación-acción. Funciones, fundamentos e instrumentos*. Barcelona. Laertes.
- KEMMIS, S. y McTAGGART, R. (1988): *Cómo planificar la Investigación-acción*. Barcelona. Laertes.
- LÓPEZ GORRIZ, I. (1991): Exigencias educativas del modelo de investigación-acción. *Enseñanza*, 8, 130-142.

DIAGNÓSTICO-EVALUACIÓN: COOPERACIÓN Y COMPLEMENTARIEDAD

Santiago Castillo Arredondo
Profesor Titular. UNED. Madrid

1. INTRODUCCIÓN

Una contestación inmediata, a bote pronto, a la pregunta que formula este Simposio debe venir de la autoridad lingüística y epistemológica que aportan Diccionarios como el *Diccionario de la Lengua* de la Real Academia Española, el *Gran Diccionario de Sinónimos y Antónimos*, de Espasa-Calpe; o *Diccionario de Ciencias de la Educación*, de Anaya, y *Diccionario de las Ciencias de la Educación*, de Santillana/Diagonal, entre otros: se diferencian los términos de *Evaluación* y *Diagnóstico*, porque etimológicamente y conceptualmente son diferentes.

Se trata de conceptos —diagnóstico y evaluación— de utilidad y aplicación universal. No son exclusividad de ninguna disciplina, área de conocimiento, o actividad profesional. En este sentido, sería un error seguir asimilando o confundiendo, el diagnóstico con el ámbito clínico o médico, o con la orientación educativa y, ahora con la evaluación. En la actividad laboral, o social encontramos el *diagnóstico* en el taller de reparación de vehículos. Por dicha sección de «diagnosis» o de «diagnóstico» pasa el vehículo **antes de** que el mecánico emprenda la reparación que corresponda. El médico nos pide realizar análisis clínicos específicos para poder diagnosticar una dolencia, **antes de** proceder a la necesaria intervención quirúrgica; y un responsable social realiza un diagnóstico de necesidades en una determinada área social **antes de** diseñar el programa que de respuesta a sus necesidades. Luego vendrá el proceso evaluativo que, sobre información producida, de principio o fin, en un programa, intervención o actuación de cualquier tipo social o laboral, valorará críticamente el acierto o adecuación, tanto en su planificación previa, como en el desarrollo seguido y en los resultados conseguidos.

2. APLICACIÓN DEL DIAGNÓSTICO EN EL ÁMBITO DIDÁCTICO

En el ámbito escolar, y de la educación en general, el diagnóstico educativo, en cualquiera de sus modalidades —diagnóstico pedagógico, diagnóstico didáctico, diagnóstico psicológico, etc.—, cobra su específico sentido. El sujeto-alumno que llega al aula al inicio del curso carece de historial personal para el profesor y para el Centro en muchas ocasiones; tan sólo se conoce de él los datos de inscripción y poco más. Esa carencia ha de ser atendida por un proceso diagnóstico que aporte la necesaria información para conseguir el mejor conocimiento del alumno. El diagnóstico educativo ha de proporcionar al profesor datos precisos sobre la realidad individual de cada alumno: información referente a su ambiente familiar, a su desarrollo madurativo y características personales, a su contexto social, a sus posibilidades intelectuales, a sus posibles carencias o dificultades, etc. Sólo sobre la base de un

conocimiento profundo y completo de cada uno de sus alumnos, el profesor podrá diseñar la actividad educativa del aula, pudiendo ajustar y adaptar las actividades a las peculiaridades de cada uno de ellos.

El dominio en la práctica del diagnóstico es una de las competencias que conforma la responsabilidad profesional del profesor en todos los niveles de la enseñanza. El ejercicio de la docencia requiere una preparación actualizada respecto a la especificidad de la etapa y a la problemática diversa que se le puede presentar. Los planteamientos didácticos de la reforma educativa requieren las aportaciones del diagnóstico educativo. El proceso diagnóstico forma parte del diseño curricular y se complementa con el proceso de evaluación. Su contribución será imprescindible en situaciones como las adaptaciones curriculares, la apreciación de los ritmos de aprendizaje, o la atención a niños con necesidades educativas especiales, entre otros.

3. APORTACIÓN DE LA EVALUACIÓN

La evaluación recibe un tratamiento de excepción en el nuevo ordenamiento educativo. Su importancia sobrepasa la incidencia que tiene en el desarrollo de la actividad curricular en el aula y las consecuentes repercusiones sociales. En el Diseño Curricular la evaluación tiene justificada su importante aportación por la responsabilidad que se le asigna respecto a la calidad de la enseñanza y una mejor eficacia de todos los elementos que intervienen en la actividad educativa.

La mejora de la calidad de la educación, la optimización de cuanto acontece en el proceso de enseñanza-aprendizaje, y, en definitiva, el perfeccionamiento de todo el sistema educativo, es la preocupación constante de la reforma educativa y de sus responsables. No hay mejora sino hay un conocimiento real de la situación educativa; y ello no es posible si no se cuenta con una evaluación eficaz que posibilite la oportuna toma de decisiones.

4. COOPERACIÓN Y COMPLEMENTARIEDAD

Diagnóstico y evaluación suponen el principio y el fin de todo proceso didáctico: uno lo abre y otra lo cierra. Pero a lo largo de todo este proceso educativo «cooperan» y se «complementan» estrechamente, reafirmando la unidad secuenciada del proceso didáctico. El diagnóstico se comporta como proceso dinámico encaminado a explorar y analizar, descubrir y describir (=conocer más y mejor) la actividad escolar que se intenta desarrollar; determinar el marco diferencial de dicha actuación; fijar los objetivos y las metas —predicción y pronóstico—; detectar anomalías y dificultades —etiología—. La evaluación, a su vez, supone el seguimiento crítico del proceso que permite recoger información con el objeto de poder establecer un juicio valorativo con intencionalidad perfectiva y sancionadora de la actuación didáctica, ya sea provisionalmente en sus comienzos, formativamente durante su ejecución, o terminalmente al final de su realización.

El proceso didáctico se puede sintetizar en cinco fases. Las cuatro primeras —diagnóstico, planificación, ejecución y evaluación— constituyen el ciclo básico de todo un proceso que se repite a su vez ciclicamente. Están concatenadas entre sí, y orientadas a un nuevo diagnóstico, que puede modificar el anterior (lo que constituye la quinta fase), y, en su caso, llevar a una nueva planificación y a nuevas formas de intervención, repitiéndose ciclicamente todo el esquema.

El diagnóstico y la evaluación, cada cual con su especificidad, constituyen un conjunto de indagaciones procesuales y sistemáticas utilizadas para conocer el desarrollo de la actuación didáctica con la intención de proponer sugerencias y pautas perfectivas en todas y cada uno de los elementos integrantes en el proceso de enseñanza-aprendizaje.

Son dos procesos que interactúan constantemente y que utilizan algunos elementos comunes. Así, el análisis de los resultados obtenidos en la evaluación del rendimiento de un alumno permite diagnós-

ticar su desarrollo posterior y, al mismo tiempo, los datos aportados por un diagnóstico facilitan la interpretación de los resultados de la evaluación siguiente. Sin embargo, sus fines difieren, ya que, así como la evaluación sirve para determinar en qué grado se lograron los objetivos propuestos, mediante el diagnóstico se descubren y exploran dificultades que pueden interferir en el proceso enseñanza-aprendizaje. Conocidas las causas, el profesor estará en condiciones de ayudar a cada alumno para que alcance el máximo nivel que sus capacidades le permiten.

El profesor de aula tiene que aunar, de alguna forma, en su práctica docente, las funciones diagnósticas y evaluadoras, agregando, incluso, las aportaciones de los especialistas, ya que desde ámbitos diferenciales interactúan y se benefician mutuamente.

5. SIMILITUD Y DIFERENCIA

«Diagnóstico» y «Evaluación» son dos conceptos etimológica y epistemológicamente diferenciados en la literatura pedagógica. En la actualidad, como consecuencia de su confluencia en el mismo campo de intereses e intervenciones educativas, se han aproximado también sus campos conceptuales hasta llegar a un cierto «confusionismo» terminológico. Se utiliza indistintamente un término u otro para referirse a situaciones o hechos educativos, como si fuesen conceptos sinónimos o intercambiables.

Una mayor precisión y rigor conceptual en el uso del lenguaje específico pedagógico y didáctico, nos debe llevar a ver en el diagnóstico y en la evaluación, más la «cooperación» y la «complementariedad» de dos procesos, que la simplificación y la imprecisión que lleva a una sinonimia académica y lingüísticamente incorrecta. En términos generales la evaluación se dirige fundamentalmente a la mejora de la calidad educativa, evaluando los distintos aspectos del contexto educativo para tomar decisiones que posibiliten dicha mejora. El diagnóstico por su parte, tiene como finalidad descubrir y describir, predecir y, en su caso, explicar el comportamiento del individuo, grupo o situación, estudiando todos los factores intervinientes (personales, interactivos, contextuales). El diagnóstico educativo posibilita el análisis de necesidades, proporciona información sobre estrategias de intervención, ayuda con la información que aporta a facilitar la toma de decisiones que se concretiza en adecuadas estrategias de intervención y/o programas de tratamiento, o en garantizar un adecuado diseño y ejecución didáctica.

Es evidente que diagnóstico y evaluación comparten aspectos comunes, sobre todo, en su tarea informativa. Participan de técnicas e instrumentos comunes. Por el «momento», la «intencionalidad» y el «tipo» de información tienen claras diferencias. El diagnóstico proporciona una información «anticipada y previsor» (la que a algunos les induce a confundir con la «evaluación diagnóstica» y de ahí, por extensión, sumir al diagnóstico dentro de la evaluación, en una concepción simplificadora), en el momento previo (= antes de) como «punto de partida» precedente a toda actuación. Es un tipo de información específica, inicialmente oculta, o de difícil acceso, que tras largo proceso desde lo aparente y general va profundizando hasta llegar al dato oculto y puntual en que nos pone de manifiesto la raíz causal y explicativa de un hecho o situación. El diagnóstico proporciona una información relacionada, precisa y específica, como garantía de acierto en las decisiones posteriores.

La evaluación, por su parte, va acumulando todo tipo de información; también la que le proporciona el diagnóstico, tal como se va produciendo en todos los momentos de una actividad educativa, de principio a fin. Esta información se recoge «después de», o como «punto de llegada» en la conquista de pequeñas metas —formativamente— o en los resultados de logros finales, con el objeto de otorgarle una valoración determinada. La evaluación requiere una información amplia, diversa, observable y «palpable», secuenciada, que fundamente un adecuado enjuiciamiento valorativo, desde el cual poder tomar la decisión pertinente.

Diagnóstico y evaluación van más allá de su propio concepto; son un proceso, una forma de proceder que requiere una actitud de búsqueda y de clarificación, igualmente diferenciadas.

LA EVALUACIÓN DE LA LECTURA: VARIABLES IMPLICADAS EN EL RECONOCIMIENTO DE PALABRAS

Sylvia Defior, Fernando Justicia y Francisco J. Martos
Universidad de Granada

Desde una perspectiva educativa, la lectura es uno de los aprendizajes básicos que han de realizar todos los alumnos. Las razones de esta obligatoriedad se derivan del doble objetivo que se pretende lograr con la lectura: debe ser un aprendizaje en sí mismo —*aprender a leer*— y, una vez conseguido, se convierte en un instrumento básico para la realización de nuevos aprendizajes —*leer para aprender*—. Por otra parte, la importancia de la lectura para el desarrollo cognitivo, afectivo y social de los individuos ha sido resaltada por diversos especialistas (Stanovich, 1986). La adquisición de la lectura, sin embargo, resulta problemática para muchos niños y adultos. Además, las dificultades de aprendizaje, en general, también se asocian en un 60% de los casos con problemas en el dominio de la lectura (Westman, Ownby y Smith, 1987).

El presente trabajo se interesa por aspectos relacionados con la evaluación de la lectura. La evaluación debe ser una característica esencial de todo proceso de enseñanza-aprendizaje y cobra mayor relevancia cuando tratamos de evaluar una habilidad tan compleja como la lectura, tanto en su desarrollo normal como en el caso de los niños que presentan dificultades en el aprendizaje lector. En nuestro país los instrumentos disponibles para evaluar la lectura no suelen tener en cuenta los actuales modelos explicativos de los procesos cognitivos implicados en la misma. La elaboración de pruebas de lectura fundamentadas en estos hallazgos es una cuestión pendiente en nuestro ámbito.

Los modelos teóricos actuales conciben la lectura como una habilidad cognitiva compleja en la que intervienen diferentes procesos que al interactuar producen finalmente la comprensión del texto. Básicamente, podemos agrupar los procesos que intervienen en la lectura en dos grandes grupos: los procesos de reconocimiento de palabras, llamados de «bajo nivel» y los procesos de «alto nivel» más relacionados con la comprensión lectora. En cada uno de ellos, a su vez, se dan una gran diversidad de subprocesos.

Los procesos de reconocimiento incluyen aspectos perceptivos, dedicados al análisis de los rasgos visuales y a la identificación de las letras, aspectos fonológicos de transformación de esos patrones visuales en sonidos, y también los llamados procesos de acceso léxico que son los implicados en la obtención del significado de las palabras. Pero normalmente no leemos palabras aisladas sino integradas en frases, párrafos y textos, con el fin de extraer la información que contienen. En este segundo gran componente, dedicado fundamentalmente a la comprensión, intervienen procesos sintácticos, inferenciales, los conocimientos previos del lector y su organización, etc. Todo ello hace muy compleja su evaluación.

En cualquier caso, las pruebas de lectura deberían ajustarse a la evaluación de todos estos componentes asociados a los procesos que la investigación en los últimos años ha puesto de manifiesto.

En esta comunicación, sin embargo, nos interesa tan sólo resaltar la importancia de las variables que influyen en los procesos de reconocimiento de palabras. Situándonos a este nivel, los modelos proponen la existencia de una doble ruta: la léxica, directa o visual y la no léxica, indirecta o fonológica. Después de un amplio debate entre los defensores de una u otra en los años setenta, en la actualidad existe un gran consenso entre los investigadores acerca de la presencia y disponibilidad de ambas vías en los lectores hábiles, si bien con predominio de la primera. Aunque, en las fases iniciales de la adquisición de esta habilidad la ruta fonológica sea la predominante.

La investigación actual necesita, por lo tanto, identificar y explorar el comportamiento de las variables que influyen en la lectura de las palabras. Nos referimos a las variables de frecuencia léxica, tipo de palabra, longitud, estructura silábica, imaginabilidad, polisemia; etc. Así, por ejemplo, las palabras frecuentes producirían menos errores de lectura que las menos frecuentes. Existen numerosas pruebas de la importancia de estas variables (Segui, Frauenfelder, Laine y Mehler, 1987, en lengua inglesa; en nuestra lengua García-Albea, Sánchez-Casas y del Viso, 1982; Valle, 1989).

Todos estos trabajos hacen hincapié en la importancia de las variables mencionadas aunque sólo exploren su influencia de forma parcial. Nosotros estamos llevando a cabo un análisis exploratorio, que situado en el contexto de un diseño de investigación más amplio, tiene como objetivo dar a conocer la influencia de todas las variables que, a nivel léxico, subléxico, contextuales y de sujeto, pueden influir en la lectura del español. A pesar de que en nuestro trabajo, hasta el momento, sólo se han analizado los datos muy globalmente, existen evidencias de que los factores mencionados afectan significativamente al número de errores que cometen los sujetos en un test de palabras aisladas (ver Tabla I).

En concreto nuestra prueba se componía de 306 palabras en las que se estudió la influencia de las variables que se relacionan:

Tipo de palabra: Se utilizaron 193 palabras, 59 pseudopalabras (sin representación léxica) y 54 no palabras (sin representación léxica y con combinaciones ortográficas inexistentes en español aunque pronunciables). Las pseudopalabras y no palabras se construyeron cambiando una letra de las palabras; un ejemplo de cada una de ellas sería: playa, blaya, mlaya.

Con respecto a la **función lingüística** de las palabras, 175 eran palabras-contenido (Nombre, pronombre, adjetivo, verbo, adverbio) y 18 eran palabras-función (Artículo, conjunción, preposición). Las restantes 113 no tenían ninguna función al ser pseudo o no-palabras.

Frecuencia léxica: Las 193 palabras se seleccionaron de acuerdo con el diccionario de frecuencias de Justicia (1990): 59 de frecuencia alta, 54 de frecuencia media y 80 de frecuencia baja. Las restantes 113 al ser pseudo o no-palabras tenían frecuencia cero.

Longitud: En el total de la prueba, 53 palabras eran monosílabas, 76 bisílabas, 91 trisílabas y 86 eran de cuatro o más sílabas.

Estructura silábica: En la composición de las palabras se tuvieron en cuenta las siete estructuras silábicas más frecuentes del español, exceptuando las sílabas que coinciden con cada una de las vocales (39 CV, 43 VC, 48 CVC, 42 CCV, 45 CVVC, 39 CVV, 41 CCVC).

Posición silábica: Se tuvo en cuenta la posición de la sílaba manipulada. Así, 171 palabras tenían la sílaba en primera posición, 64 palabras la tenían en segunda posición, 49 en tercera y 22 en cuarta o quinta posición.

Frecuencia silábica: se ha puesto de manifiesto que esta variable puede tan influyente como la frecuencia léxica. En nuestro estudio se quiere analizar también su influencia en los errores de lectura.

Igualmente se tuvieron en cuenta una serie de variables referentes a las características de los sujetos tales como **sexo**, **nivel lector** y **clase social** así como variables de tipo contextual como **curso**,

TABLA 1

MEDIAS Y DESVIACIONES TÍPICAS DE LOS ERRORES COMETIDOS EN LA LECTURA DE LAS PALABRAS EN FUNCIÓN DE CADA UNO DE LOS NIVELES¹ DE LAS VARIABLES CVF (CONTE-NIDO VERSUS FUNCIÓN), TIP (TIPO DE PALABRA), FRE (FRECUENCIA LÉXICA), LON (LON-GITUD), PSI (POSICIÓN DE LA SÍLABA) Y ESI (ESTRUCTURA DE LA SÍLABA). TODAS LAS COMPARACIONES ENTRE TODOS LOS NIVELES DE CADA UNA DE LAS VARIABLES RESUL-TARON SER SIGNIFICATIVAS EXCEPTO EN LA VARIABLE ESI EN LAS QUE NO LO ERAN LAS COMPARACIONES ENTRE LOS NIVELES 2 Y 3 Y ENTRE LOS NIVELES 4 Y 6.

NIV VAR		1	2	3	4	5	6	7
CvF	M	45.04	25.58	13.57				
	SD	19.89	23.49	11.06				
TIP	M	24.46	51.56	39.07				
	SD	22.14	21.03	19.78				
FRE	M	45.04	15.82	23.59	31.43			
	SD	19.89	19.66	24.38	23.78			
LON	M	15.15	28.43	34.37	43.25			
	SD	14.26	20.31	23.79	24.79			
PSI	M	27.72	28.95	39.29	45.65			
	SD	19.31	19.88	23.72	24.82			
ESI	M	25.13	29.42	29.44	32.59	34.75	32.22	37.81
	SD	18.11	19.35	21.28	21.40	23.07	23.67	22.17

¹ Niveles:

CvF: 1 (No palabras y Pseudopalabras), 2 (Palabras-contenido), 3 (Palabras-función).

TIP: 1 (Palabras), 2 (No palabras), 3 (Pseudopalabras).

FRE: 1 (No palabras y Pseudopalabras), 2 (Frecuencia Alta), 3 (Frecuencia Media), 4 (Frecuencia Baja).

LON: 1 (Monosílabas), 2 (Bisílabas), 3 (Trisílabas), 4 (Cuatro o más sílabas).

PSI: 1 (Primera), 2 (Segunda), 3 (Tercera), 4 (Cuarta o quinta).

ESI: 1 (CV), 2 (VC), 3 (CVC), 4 (CCV), 5 (CVVC), 6 (CVV), 7 (CCVC).

tipo de colegio (privado o público) y método de enseñanza de lectura (global, fonético y mixto). Estas variables no van a ser presentadas en este trabajo.

Como variables dependientes se utilizaron el número de errores y el tipo de error cometido: Silabeo, repetición, sustitución, adición, omisión, rotación, inversión, lexicalización y error de acentuación.

Se aplicó esta prueba a 140 sujetos de enseñanza primaria, veinte por cada curso de 1º a 6º de E.G.B. más 4 retrasados lectores por cada curso de 2º a 6º. Consideramos como retrasados a aquellos

niños cuyo nivel lector era de dos o más cursos por debajo del correspondiente a su edad cronológica. Los sujetos procedían de cuatro colegios, dos públicos y dos privados. La selección de los sujetos se hizo al azar, con la restricción de que hubiese el mismo número de niños que de niñas. También se tuvo en cuenta que los grupos estuviesen equilibrados en cuanto a la clase social. Las respuestas de los niños se anotaban en su hoja correspondiente, al mismo tiempo que se grabaron magnetofónicamente.

En el nivel actual de desarrollo de nuestro trabajo ya han sido analizados los datos que demuestran la influencia de algunas de las variables. En la Tabla I se exponen estos resultados. Puede comprobarse al constatar la tabla, que todas y cada una de las variables consideradas producen efectos significativos en cuanto al número de errores cometido por los sujetos.

Como conclusión queremos recalcar la importancia y la necesidad de que estas variables sean tenidas en cuenta en el diseño de instrumentos de evaluación de la madurez lectora. Al hacerlo no sólo se conseguiría un diagnóstico más válido y fiable sino que además nos permitiría determinar con precisión el tipo de errores que comete el sujeto y las variables que lo influyen. Esto redundaría en una intervención educativa más eficaz.

REFERENCIAS

- GARCÍA-ALBEA, J. E.; SÁNCHEZ CASAS, R. y del VISO, S. (1982): Efecto de la frecuencia de uso en el reconocimiento de palabras. *Investigaciones Psicológicas*, 1, 24-63.
- SEGUI, J., FRAUENFELDER, U. H., LAINE, C. y MEHLER, J. (1987): The word frequency effect for open —and closed— class items. *Cognitive Neuropsychology*, 4, 33-44.
- STANOVICH, K. (1986): Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, 21, 360-407.
- VALLE, F. (1989): Errores en lectura y escritura. Un modelo dual. *Cognitiva*, 2, (1), 35-63.
- WESTMAN, J. C., OWNBY, R. L. y SMITH, S. (1987): An analysis of 180 children referred to an University Hospital Learning Disabilities Service. *Child Psychology and Human Development*, 17, 275-282.

LA FIGURA DE LAS LETRAS CURSIVAS

Ignacio Alfaro*, Francisco Secadas**

* Universitat de València

** Universidad Autónoma de Madrid

ANTECEDENTES Y PLANTEAMIENTO DEL PROBLEMA

Después de una serie de estudios sobre la escritura se ha podido diseñar, y experimentar durante cuatro cursos académicos con niños valencianos y barceloneses, una nueva metodología de enseñanza de la escritura (Secadas y otros, 1993). La hipótesis del sistema de enseñanza gira en torno a un planteamiento inicial sencillo: *Las letras tienen una figura; existen unos grupos de las letras según su figura y, evolutivamente, se puede establecer una sucesión acorde con el grado creciente de dificultad inherente a las figuras*. El planteamiento didáctico sigue los principios psico-evolutivos investigados. La enseñanza de la escritura se establece a los cinco años y se sustenta en un tipo de letra, la *escrip*, que difiere en algunos rasgos de la *script* de Dotrens. Esta letra ha sido diseñada al efecto y está disponible para su edición por computador. Asimismo, se acompaña un análisis evolutivo de los diferentes grupos de letras, de forma que el profesor puede comparar el rendimiento y dominio de la escritura de sus alumnos en relación a su grupo de edad. En caso de que este dominio sea insuficiente, podrá igualmente acudir al análisis de errores de los grafismos o buscar el momento evolutivo en el que se encuentra la habilidad de un alumno concreto para, a partir de las automatizaciones conseguidas, proponerle una secuencia encaminada a la consecución de nuevas habilidades.

Una vez que los niños han aprendido a escribir con este método y han automatizado, en particular, el dominio de la forma de las letras *escrip*, se hace necesario abordar un proceso de *transición* hacia la *letra cursiva y ligada*. Este proceso se hace mucho más sencillo si confluyen dos factores: una automatización previa de la letra *escrip* y una propuesta de transición acorde con el grado progresivo de dificultad del tipo de letra cursiva. Nuestro proceso de enseñanza coincide en algunos aspectos con otros métodos en los que se propone inicialmente un tipo de letra imprentoide (D'Nealian Manuscript), para facilitar la transición a la cursiva (Scott, Foresman & Company, 1982), salvo en un importante detalle: que al hacer inclinados los trazos ortogonales de la *escrip* se convierten en oblicuos, añadiendo con ello un factor innecesario de dificultad, según se desprender de nuestros trabajos.

En los estudios experimentales realizados antes de la publicación del método, se ha contrastado no sólo el nivel de ejecución de los grafismos (*escrip* y *cursivo*) antes y después de la enseñanza de la escritura, sino también su grado de dominio y tiempo empleado en la copia en relación al tipo de letra utilizada por el sistema de enseñanza adoptado en cada muestra. En este sentido, podemos adelantar que el inicio de la enseñanza de la escritura con el tipo de letra *escrip* ofrece demostradas ventajas.

Los estudios en los que se basa este sistema no hubieran podido realizarse sin la aplicación a

grandes muestras de sujetos de diferentes escalas de escritura: Escala Gráfica L/S, Escala Gráfica de Edad (EGE), Escala Gráfica General (EGG) y Escala de Escritura (EE) (Secadas y otros, 1984, Rodríguez, M. T., 1986 y Secadas y otros, 1993). Para la realización de los contrastes mencionados, se hizo necesario la construcción de unos instrumentos de medida de las letras *cursivas*, semejantes a los utilizados para la letra *escrit*. En este trabajo presentamos los resultados y análisis realizados en el departamento MIDE de Valencia bajo la dirección del primer firmante, para la detección de las formas arquetípicas de las letras de tipo cursivo y sus correspondientes *grupos de letras* en orden a la elaboración de una ESCALA GRÁFICA CURSIVA L/S derivada de los análisis anteriores.

HIPÓTESIS Y OBJETIVOS

Partimos de la hipótesis de que la enseñanza de la escritura debe iniciarse con la letra *escrit* en los 5 años, y posteriormente introducir la cursiva de una forma paulatina y acorde con el grado progresivo de dificultad de cada grupo de letras. La calidad de la escritura dependerá de múltiples factores: forma, tamaño, espaciamiento, alineamiento, nitidez, etc. pero, sobre todo, de que el método de enseñanza propuesto vaya paralelo al desarrollo madurativo del sujeto. En este sentido, uno de los pilares básicos de la calidad de la escritura pasa por el tipo de letra, la forma de las letras y la capacidad del niño para realizar ciertos trazos. No existe una calidad final de la escritura sino una relativa a la habilidad gráfica del niño en cada momento. Nos parece arbitrario que Graham y Madan (1981) afirmen que: «la mayoría de errores son malformaciones y únicamente unos pocos símbolos —*a, e, r, t*— acumulan alrededor del 50 por ciento de todas las malformaciones de las letras, en cualquier nivel educativo». En todo caso, este análisis dependerá, además, de otros factores endógenos del sujeto, del tipo de letra utilizado y del momento evolutivo del niño. En ciertas edades no es posible esperar una buena realización de algunos modelos de letras. Por ello, si hemos diferenciado anteriormente unos grupos de letras y una secuenciación madurativa para la letra *escrit*, también se puede hacer otro tanto con la letra *cursiva*, de forma que la transición de una a otra letra sea gradual y se propongan criterios concordados de enseñanza.

Básicamente, los estudios revisados establecen grupos de letras cursivas y secuencias para su aprendizaje en muchos de los casos completamente discordantes entre sí o inconsistentes. La mayoría de trabajos establecen intuitivamente la afinidad de los grupos de letras por su forma. Es un buen principio, si se tienen en cuenta los trabajos de E. J. Gibson y colaboradores, que sostenían que la percepción visual antecede al juego motriz con el que interactúa. Por otra parte, en bastantes de estos estudios los grupos de letras presentados responden al sistema de enseñanza en que se desarrollan, con lo que su utilización es limitada. Otros trabajos ofrecen una secuencia y grupos de letras cursivas en base a dos criterios: su forma y la similaridad de la direccionalidad del movimiento. Y todavía otros añaden criterios nuevos para su agrupación, basados en estudios optométricos y en la coordinación visual y táctil. Curiosamente, cuando se mezclan tantos criterios de una forma globalizadora sin una hipótesis psico-evolutiva estricta que guíe el sistema de enseñanza, aparecen en un mismo autor diferentes propuestas de secuencias y agrupaciones de letras que pueden variar con dos años de diferencia (Getman, 1983 y 1985). En nuestras experimentaciones, hemos preferido utilizar un sistema de medición paralelo en ambos dos tipos de letra (*escrit* y *cursiva*) y que sirviera de instrumento para analizar la evolución gráfica en cada uno de los sistemas de enseñanza.

En relación al **primer objetivo mencionado**, la detección de los arquetipos básicos de la letra cursiva, el análisis dimensional diferencia por su forma seis grupos de letras que pueden dar la clave para la transición de un estilo a otro, objeto que será de otro trabajo y autor. Tocante al **segundo objetivo**, se presenta la ESCALA GRÁFICA CURSIVA L/S. Su aplicación a muestras representativas de niños de 4 a 7 años nos ha servido para detectar el grado de madurez de los niños para la realización de ciertos grafismos.

BÚSQUEDA DEL ALFABETO CURSIVO

La primera tarea, antes de analizar la forma de las letras, fue buscar un alfabeto que fuera representativo de lo que se entiende normalmente por letra cursiva o ligada. Para ello, se revisaron todos los alfabetos de las editoriales implantadas en nuestro mercado. Como primera observación, podemos afirmar que no existen criterios homogéneos en los alfabetos cursivos. Frecuentemente, se entremezclan en un mismo alfabeto tipos de letra cursivos con tipos «mixtos», tanto en mayúsculas como en minúsculas y, más aún, si alguna letra cursiva original presenta grandes dificultades se sustituye directamente por la correspondiente de tipo «script» aunque sea imposible su «ligamiento» con el resto o su forma no sea la típicamente cursiva. El proceso de selección de las letras cursivas fue acordado por jueces y el alfabeto puede observarse en la *tabla 1*.

METODOLOGÍA PARA EL ANÁLISIS DE LA FIGURA DE LAS LETRAS

El instrumento utilizado para analizar las letras cursivas fue el *análisis dimensional* de Secadas, que permite decantar grupos o dimensiones según los grados de afinidad establecidos entre las letras del alfabeto, mayúsculas, minúsculas y los números. Cada par de letras fue sometido a una estimación por jueces acerca de su semejanza o afinidad en una escala de 7 intervalos (de +3 a -3), utilizando criterios como: semejanza aparente de sus formas o mala forma en términos gestálticos, número y tipo de trazos comunes, identidad posicional de sus elementos, formas con tendencia al cierre, rupturas y modificaciones de las formas, etc.

El conjunto de las estimaciones de los jueces fue promediado con el fin de construir una matriz única de afinidades entre todas las variables (letras y números) a la que pudiera aplicarse el *análisis dimensional* (programa *Anadim2* de J. Sanmartín, 1985a), que obtuvo una matriz dimensional de primer grado de 29 dimensiones. A este conjunto dimensional resultante se le aplicó el teorema de Bayes en sucesivas iteraciones (*Bayes 2* de J. Sanmartín 1985c) con el fin de hallar las probabilidades de cada una de las saturaciones dimensionales, condicionadas a su fila y columna correspondiente, solución parecida a la rotación factorial que ajusta las saturaciones dimensionales de las variables a sus ejes.

A partir de la mayor o menor afinidad (elementos comunes) entre las 29 primeras dimensiones se confeccionó una segunda matriz (*Matcova2* de J. Sanmartín, 1985b) de la que se obtuvieron nuevas dimensiones de segundo grado (15 dimensiones) y, con procesos sucesivos semejantes, se llegó a un análisis de tercer grado (6 dimensiones). Los tres grados de análisis han permitido ir reduciendo el rango matricial y agrupando las dimensiones en conjuntos progresivamente más sintéticos, que facilitan la comprensión de los esquemas figurales de las letras, sin perder por ello las referencias originales de las primeras variables, las letras.

LA FIGURA DE LAS LETRAS: RESULTADOS DIMENSIONALES

La *tabla 1* sintetiza los resultados dimensionales obtenidos. Las 29 dimensiones de primer grado agrupan las letras con características comunes a cada dimensión. En el segundo grado se obtienen 15 dimensiones de las que se pueden extraer las características de los **esquemas radicales** de las formas cursivas comunes a sus letras. Las dimensiones de tercer grado reflejan las seis estructuras básicas que configuran el alfabeto cursivo y que comentamos seguidamente.

Dimensión 1: Redondas (versus aristas abiertas). La primera dimensión (abscisa) viene definida por su polo positivo. El rasgo básico de las figuras de las letras que la conforman son las formas redondas, tanto completas como incompletas, que encierran partes circulares o curvas, cuyo elemento

Dimensiones.

3º Grado	2º Grado		1º Grado	
D1 O	D1 O	D1 a o d c	D8 q 9	D17 O O O
	D3 e	D4 e e e f	R8 6	
	D9 B	D14 B D P R	R9 8	
D2 f	D2 l	D2 b l ll	D3 h k h ch	D15 L Ll
	R1 y	D5 f g j y g j y		
D3 i	D4 i	D6 i i	R2 t	
	D6 s	R1 r	D9 s j	
D4 n	D5 n	D7 m n ñ p	D10 u v U V	
	R2 w	D11 w W		
D5 x	D7 x	D12 x h x	R3 A	
	R3 k	R5 k		
	R4 n	D16 M N X		
D6 Z	D8 Z	D13 z Z E F	R4 F	
	R5 L	R6 L		
R1	D10 Z	D18 Z 3	R7 5	

TABLA I
SÍNTESIS DIMENSIONAL DEL ANÁLISIS DE LAS LETRAS CURSIVAS

común es el círculo. Su polo negativo reafirma esta característica al tratarse de formas abiertas con aristas.

Dimensión 2: Verticales con volutas. En este grupo, no se interrumpe el trazado de la letra. Los arquetipos son fundamentalmente ortogonales y los oblicuos son substituidos por lazos y volutas, en vez de por aristas o líneas onduladas, con el fin de enlazar las letras. Cuando se puede obtener un trazo curvo se prefiere al oblicuo o interrumpido. Esta consideración es clave para la transformación de la letra *escrita a cursiva*, ya que se considera que los arquetipos oblicuos son los más difíciles.

Dimensión 3: Líneas descendentes. Las figuras de este sector representan los «palos» verticales, aunque se pueden diferenciar dos variantes: a) los palos oblicuos con descensos ondulados, cuya forma es simple aunque requiere una precisión en la parte inferior del trazado, elemento fundamental en la formación de la letra (dimensiones: R1 y D9); y b) los palos simples con tramos, tildes y puntos, interrumpiéndose el trazado de la letra en algún caso (dimensiones: D6 y R2).

Dimensión 4: Paralelas unidas con curvas. El arquetipo básico está formado por trazos verticales cuya unión, ya sea en su parte superior o inferior, se realiza a través de una ondulación de la línea, pero sin voluta, con el fin de obviar los trazos oblicuos. Su resolución es mucho más simple al evitar los arquetipos oblicuos.

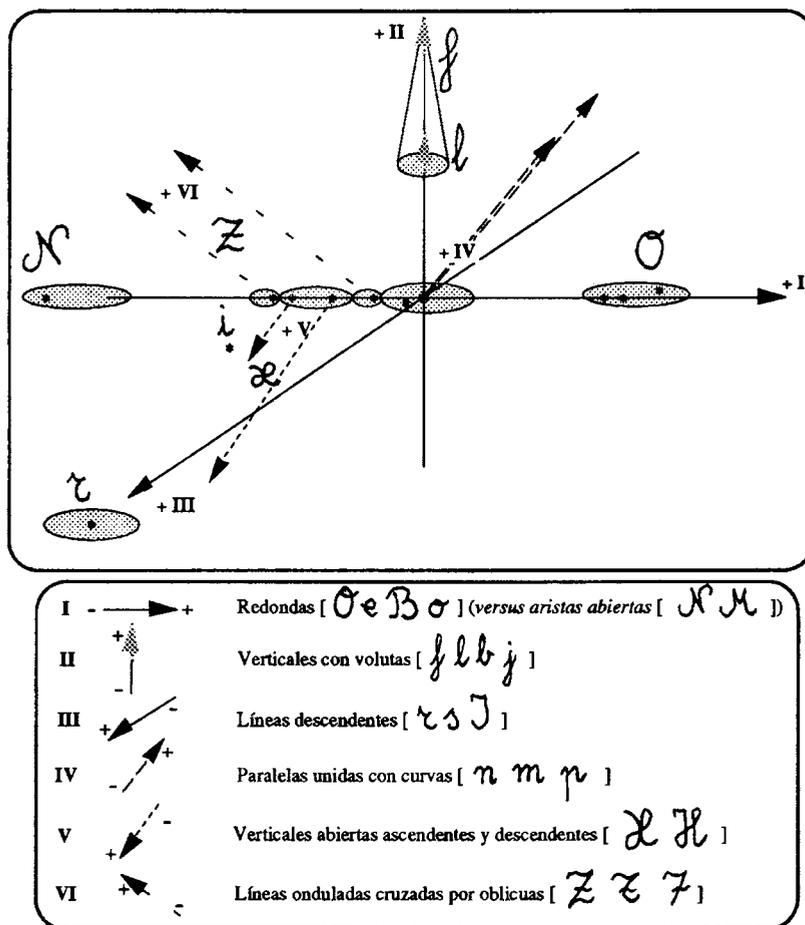


Figura 1

Representación gráfica de los esquemas figurales de la letra cursiva y su interrelación.

Dimensión 5: Verticales abiertas ascendentes y descendentes. Las figuras de este sector no tienen volutas sino aristas. Los travesaños son ondulados en vez de horizontales. Los arquetipos básicos serían *ortogonales* y *oblicuos*, fluidificando los trazos horizontales y con la presencia de trazos descendentes.

Dimensión 6: Líneas onduladas cruzadas por oblicuas. El arquetipo básico está compuesto por onduladas horizontales unidas a trazos oblicuos o que las cruzan. Se trata de evitar el trazado horizontal puro, fluidificándolo con ligeras ondulaciones.

Getman (1983) propone una secuencia de aprendizaje de las letras cursivas que considera también los patrones de movimiento del brazo. A grandes rasgos sugiere, en primer lugar los círculos, seguido de las ovaladas, las curvas abiertas y, finalmente, las diagonales. Nuestro análisis coincide globalmente con sus trabajos aunque desmiembra algo más esos cuatro grupos de letras.

LA ESCALA GRÁFICA CURSIVA L/S

La ESCALA GRÁFICA L/S para la valoración de la letra **escrit**, la formaban dos tarjetas, una de LETRAS (L) compuesta por 24 letras representativas de las dimensiones o grupos obtenidos según su semejanza de forma; y otra de SIGNOS (S), con los 24 esquemas o radicales de las mismas formas. En total, 48 figuras. Su aplicación y sistema de evaluación era relativamente sencillo: el niño copiaba cada modelo debajo, a lápiz y sin borrar, y se calificaba comparando sus copias en cada signo o letra con una escala de cinco *pentas* modelo de su edad (Secadas y otros, 1993). La fiabilidad y validez diagnóstica de la escala puede encontrarse en Secadas y otros (1993) y Alfaro (1986).

La ESCALA GRÁFICA CURSIVA L/S que se presenta a continuación sigue los mismos principios de aplicación y evaluación que la anterior. Las 24 letras, y los 24 signos, son representativos de todas las letras (mayúsculas, minúsculas y números) del alfabeto cursivo y de los rasgos arquetípicos del mismo. El hecho de proponer dos muestras, una de letras y otra de signos, responde a los estudios factoriales en los que se demostró que la habilidad de copiar una misma figura, p.e. la S, es independiente según se la considere letra o signo, lo que significa que «*escribir no es lo mismo que dibujar o copiar un modelo*» (Secadas y otros, 1993).

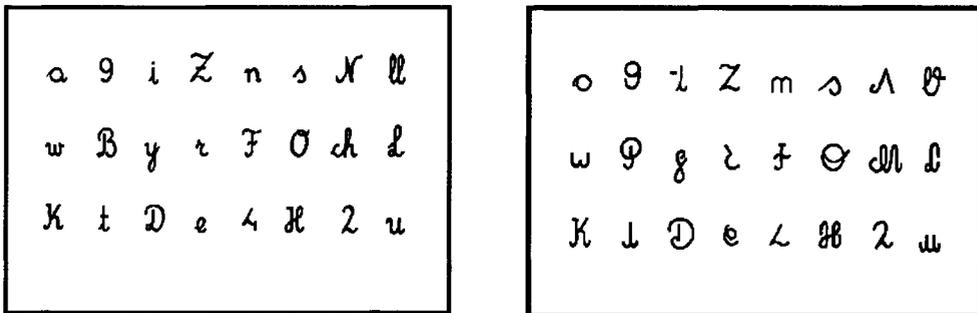


Figura 2. Escala Gráfica Cursiva L/S (letras y signos).

Los diferentes estudios relacionados con esta escala, así como los modelos *Penta* para su corrección, o el análisis de errores para cada una de las edades, se consideran, por el momento, objeto de otro trabajo.

BIBLIOGRAFÍA

- ALFARO, I. (1986): Factores gráficos en el aprendizaje de la escritura y expectativas de rendimiento. *Revista Investigación Educativa*, 4 (7), pp. 5-20.
- GETMAN, G. N. (1983): About Handwriting. *Academic Therapy*, 19 (3), pp. 139-146.
- GETMAN, G. N. (1985): Hand-Eye Coordinations. *Academic Therapy*, 20 (3), pp. 261-275.
- GRAHAM, S. Y MADAM, A. J. (1981): Teaching Letter Formation. *Academic Therapy*, 16 (4), pp. 389-396.
- RODRÍGUEZ TOMÁS, M. T. (1972): Análisis de las estructuras gráficas: bases evolutivas para una metodología de la escritura. *Tesis de Licenciatura*. Universidad de Valencia. Valencia: Inédita.
- RODRÍGUEZ TOMÁS, M. T. (1986): El aprendizaje de la escritura. Análisis psicológico y nuevo método. *Tesis Doctoral*. Universidad de Valencia. Valencia: Inédita.

- RUIZ, J. y SECADAS, F. (1984): Dificultad relativa de las letras del alfabeto. *Psicológica*, 5 (2), pp. 217-240.
- SANMARTÍN, J. (1985a): *Anadim2*. Valencia. Software: Autor.
- SANMARTÍN, J. (1985b): *Matcova2*. Valencia. Software: Autor.
- SANMARTÍN, J. (1985c): *Bayes2*. Valencia. Software: Autor.
- SECADAS, F., RODRÍGUEZ, M. T. y ALFARO, I. (1993): *Escribir es fácil*, Madrid: Ed. TEA S.A. Depósito Legal de Madrid 9128 y 9129. (2 vols.). En prensa.
- SECADAS, F., RODRÍGUEZ, M. T. y RUIZ, J. (1984): La forma de las letras. *Studia Paedagogica*, 14, pp. 99-196.
- SCOTT, FORESMAN & COMPANY (1982): *Research and information: Handwriting research and D'Nealian handwriting methods-an update*. Chicago: Autor.

DIFERENCIACIÓN DEL RIESGO EN LOS CONDUCTORES. ESTUDIO DE PERFILES

Miquel Amador Guillem
Universitat Autònoma de Barcelona

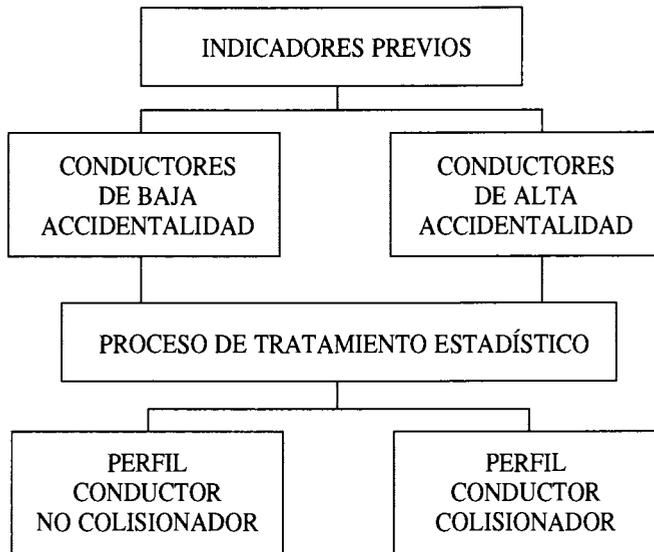
El contenido de esta comunicación resume uno de los estudios realizados en el contexto de los trabajos que sobre la FORMACIÓN DEL CONDUCTOR se llevan a cabo en el Departamento de Pedagogía Aplicada y, concretamente, en el Área de Métodos de Investigación y Diagnóstico en Educación de la U.A.B., de acuerdo al convenio suscrito con la Fundación Técnica y de Estudios del RACC (Real Automóvil Club de Cataluña).

JUSTIFICACIÓN Y PLANEAMIENTO DEL ESTUDIO

Los avances tecnológicos han venido marcando la pauta de los niveles de bienestar que pretende nuestra sociedad, siendo una de sus manifestaciones más rutilantes la del uso del automóvil. Ello ha generado una «cultura» específica no exenta de dificultades y problemas, lo cual ha supuesto la búsqueda de medidas correctoras: la mejora de las redes viarias, la descongestión del tráfico en los núcleos urbanos de mayor densidad de población y los correspondientes accesos, los catalizadores de carburante y la ausencia de plomo en los combustibles, conseguir un bajo índice de ruido, etc... Una larga lista a la que deberíamos añadir las acciones correctoras de las conductas de riesgo del automovilista, destacando la presencia de las leyes sancionadoras de las infracciones al «código de circulación», fundamentadas en la disuasión por el castigo. A pesar de lo cual no podemos olvidar los esfuerzos que instituciones de todo tipo realizan para averiguar el por qué de esas conductas de riesgo, impulsando programas de **Formación de Formadores** que, como en nuestro caso, signifiquen la intervención en lo que se conoce popularmente como «Educación Vial», dirigidos tanto a los que forman a los futuros conductores como a la mejora de la conducción de los usuarios de vehículos y al ciudadano que padece su impacto.

Poco a poco se han ido incrementado este tipo de acciones por cuanto se ha producido un fenómeno de sensibilización asociado a las informaciones que los distintos medios de comunicación presentan semanalmente sobre la accidentalidad.

Nuestro estudio, enmarcado en este contexto, tiene por objeto aproximarse al PERFIL de los conductores, tanto a los que presentan conductas de riesgo, como a la de aquellos cuya conducción no entraña peligro, denominándoles respectivamente **colisionadores** y **no colisionadores**. Esta caracterización hace que nos decidamos por un diseño de investigación fundamentado en los utilizados por MEDLEY en sus estudios sobre casos extremos de comportamiento, concretándolo como se indica en el siguiente esquema:



Los indicadores previos, aplicados mediante encuesta, nos dieron conocimiento de aquellos conductores habituales que utilizaban en mayor o menor medida los partes de las compañías aseguradoras. La información deseada se recavó en más de un millar de automovilistas, de los que obtuvimos los grupos extremos de la muestra. El cálculo se formuló como sigue:

$$\text{media de partes} = \text{n}^\circ \text{ de partes} / \text{n}^\circ \text{ años posesión carnet}$$

De esta forma se seleccionaron 153 personas que rellenaron el cuestionario utilizado como instrumento de medida, situándose finalmente la muestra productora de datos en 106 conductores.

A partir de este momento se inició el proceso de tratamiento estadístico que contemplaba dos fases. En la primera se procedió a la depuración definitiva, en función de las medias y desviaciones correspondientes a los factores estudiados. En la segunda se analizan los datos resultantes de la aplicación del cuestionario, procediendo al estudio de los correspondientes perfiles.

EL INSTRUMENTO

El cuestionario aplicado, que titulamos con el nombre de CUESTIONARIO DE EVALUACIÓN DEL PERFIL DEL CONDUCTOR, consta de dos partes. En la primera se agrupan las variables del factor general de nivel de riesgo y, en la segunda, figuran las variables correspondientes a cuatro factores de riesgo específicos: 1) factor conductor, 2) factor coche, 3) factor vías-climatología y 4) factor tránsito. Precediendo a todo ello las variables personales.

Los criterios utilizados para la elaboración del cuestionario giran en torno a tres ejes:

- la revisión bibliográfica de los estudios sobre la detección de factores que determinan las conductas de riesgo de los conductores, en los que figuran cuestionarios para este fin.
- la adecuación de los ítems seleccionados a las características propias de nuestro entorno.
- la correspondencia de las respuestas [si] y [no] con los conceptos 'correcto' e 'incorrecto'

relacionados al mismo criterio de legalidad (código de circulación), otorgándoles 0 ó 2 puntos respectivamente. A la respuesta [?] (indiferente) se la valoró con 1 punto.

Así, mediante la suma de puntuaciones, obtuvimos los resultados de cada uno de los factores estudiados. Las colas de la distribución correspondían a los casos que, convenientemente analizados, determinaron los perfiles de conductores colisionadores y no colisionadores.

Las variables originales que intervienen en el estudio se recogen en el siguiente cuadro, relacionándolas con la parte correspondiente del cuestionario:

ESTRUCTURA CUESTIONARIO	VARIABLES
DATOS PERSONALES	Personales: Sexo, Edad, Nivel de estudios, Situación laboral
PRIMERA PARTE	Factor general: Nivel de riesgo (10 ítems=10 variables)
SEGUNDA PARTE	Factor I: Riesgo Conductor (26 ítems=26 variables) Factor II: Riesgo coche (11 ítems=11 variables) Factor III: R.Vías-Climatología (13 ítems=13 variables) Factor IV: Riesgo Tránsito (6 ítems=6 variables)

Además de contemplar cada ítem como una variable, que nos proporciona información de la opinión de cada conductor sobre la acción que refleja, hemos codificado los diferentes factores como nuevas variables-síntesis del conjunto que representan.

El proceso de tratamiento estadístico se realizó utilizando el paquete informático SPSS/PC+ V 4.1, siguiendo el protocolo que le es propio.

CONCLUSIONES

De los análisis efectuados se desprenden las siguientes conclusiones:

Perfil del conductor colisionador

Para determinar el perfil del conductor colisionador, hemos establecido el rango decreciente de las proporciones de las respuestas 'incorrectas' en los ítems que componen cada uno de los factores, englobando los que tienen relación.

En la presentación de los perfiles se contemplan tres partes diferenciadas, separando los rasgos fundamentales de los que, siendo suplementarios, también deben tenerse en cuenta y, finalmente, se han globalizado las características personales más relevantes.

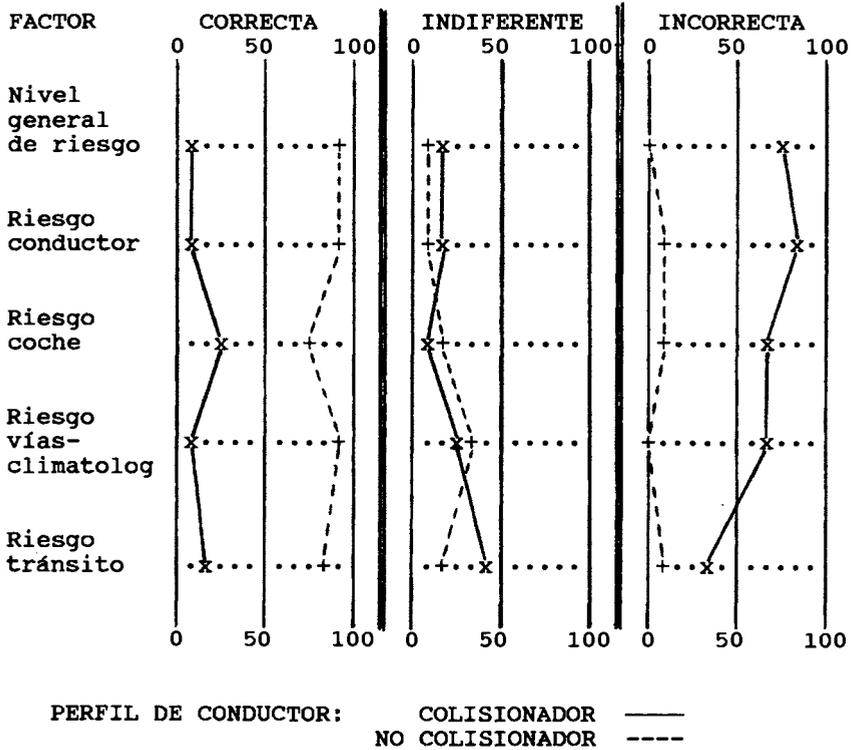
PERFIL DEL CONDUCTOR COLISIONADOR	
Fundamental	<ul style="list-style-type: none"> • Conduce tres horas seguidas sin parar • No se pone el cinturón por ciudad • Efectúa trayectos largos los fines de semana • Circula a mayor velocidad de la permitida • No está de acuerdo con el sistema de puntos • De noche conduce igual por ciudad y carretera • Adelantaría más con un GTI o similar • No prevé los golpes de viento • Ve la conducción como una aventura • Toma alcohol y conduce los fines de semana • Escoge el color del coche por gusto • Desconoce el estado de los neumáticos
Además	<ul style="list-style-type: none"> • Supera en más de 15 Km el límite de velocidad • No busca otro conductor si ha bebido • No le importa la autoescuela para aprender • No piensa en evitar el choque frontal • De noche conduce tan seguro como de día • Toma las curvas a más velocidad de lo normal • No le preocupa pasar la ITV • Se cree un conductor atrevido • Utiliza vías rápidas para poder correr más • No evita embotellamientos y le ponen nervioso • Procura adelantar cuando tiene prisa
Caract. personales	<ul style="list-style-type: none"> • Es menor de 27 años • Posee estudios Secundarios o de Diplomatura • No tiene trabajo o es esporádico • Posee el carnet desde hace 3 años como máximo • Ha recorrido menos de cien mil Km • Utiliza habitualmente el seguro, aunque no se considera responsable del siniestro • Un tercio de partes son con daños materiales y un cuarto con daños personales

Perfil del conductor no colisionador

Para este perfil se han contemplado los rangos decrecientes de las proporciones de las respuestas 'correctas'.

PERFIL DEL CONDUCTOR NO COLISIONADOR	
Fundamental	<ul style="list-style-type: none"> • Nunca conduce compitiendo con otros • No conduce si toma alcohol • Tiene cuidado conduciendo de noche por ciudad • No tendría un GTI para ir más rápido • Sabe qué hacer en situaciones arriesgadas • No avanza a menos de 150 m. de un cruce • No provoca el derrapage • Conduce a menor velocidad cuando llueve • Respeto las indicaciones de obras • Escoge el color de coche que se vea de lejos • Entra en las curvas a una velocidad prudente • Procura prever los golpes de viento • Pasa la ITV cuando le corresponde • Nunca supera la velocidad permitida • Revisa el coche para conducciones especiales • No pierde la calma en los embotellamientos • Se ajusta correctamente el cinturón • Escoge la autoescuela que mejor enseña • No se considera un conductor atrevido • No acepta un coche que no sea seguro • Conoce los puntos negros de su trayecto
Además	<ul style="list-style-type: none"> • Utilizaría el sistema de puntos • No adelantaría más si tuviera prisa • Está mentalizado para evitar el choque front • No hace trayectos largos el fin de semana • Prefiere un coche más seguro que potente • Sabe cual es el estado de sus neumáticos
Caract. personales	<ul style="list-style-type: none"> • Es Mujer de cualquier edad, Diplomada y trabajadora • Tiene el carnet de más de cinco años • Ha recorrido alrededor de cien mil Km • Utiliza el seguro una vez al año

Para finalizar, se representan gráficamente las inflexiones de los perfiles por factores en relación a la respuesta dada, utilizando la distribución media de las proporciones de ambas subpoblaciones.



BIBLIOGRAFÍA

BELTRÁN, J. (1985): Las actitudes. En J. Beltrán, Psicología Educativa, UNED, Madrid Vol. II.

FERNÁNDEZ SEARA (1978): Psicología del automovilista. Rasgos de personalidad de los conductores reincidentes. D.G.T.: Madrid.

L'HOSTE, J. (1979): Las actitudes y opiniones de los conductores. II parte. I.D.B.R.A.: Joinville-Francia.

MEDLEY, D. (1977): Review of teacher effectiveness research. En Encyclopedia of Educational Research. Mc Millan and Free Press, N.Y., pp. 1.873-1.875.

MONTANE, J.: La orientación educativa y la orientación del conductor, Educar, 8, 1985 pp. 117-124.

N.N. (1985): The Road User and Drink. Unit 6. The British Institute of Traffic Education Research: Birmingham.

N.N. (1985): The Road User. Attitudes. Unit 7. The British Institute of Traffic Education Research: Birmingham.

N.N. (1985): The Road User. Observation and Hazard Awareness. Unit 9. The British Institute of Traffic Education Research: Birmingham.

N.N. (1976): Recherche Routière. Formation de conducteurs. O.C.D.E.: París.

PRIGOGINE, Y. (1975): Manuel sur les campagnes de sécurité routière. (Guide pour la preparation, l'organisation et l'évaluation de l'efficacité. O.C.D.E. París.

SPSS Inc. (1990): SPSS/PC+ User's Guide. Chicago.

ÍNDICES CARDIOVASCULARES DURANTE PRUEBAS DE EVALUACIÓN EN EL AULA Y EN EL LABORATORIO*

Presentación A. Caballero García, Norberto Navarro Adelantado y Francisco A. García Sánchez
Facultad de Educación. Universidad de Murcia

INTRODUCCIÓN

Sabemos que la atención, el procesamiento de la información y el rendimiento pueden verse afectados por el nivel de activación fisiológica del sujeto durante la realización de una tarea, al mismo tiempo dicho nivel de activación es sensible a situaciones ambientales externas o internas al sujeto (Kerres, 1985; Redondo y Del Valle-Inclán, 1992). Una de esas situaciones, frecuente en el aula, es la prueba de evaluación o examen.

De la hipótesis de Lacey (Lacey, 1967; Lacey y Lacey, 1970) acerca de la relación entre la activación cardiovascular y la activación cerebral, se deduce que la atención a estímulos externos se acompaña de desaceleración de la frecuencia cardíaca (FC) y disminución de la presión sanguínea, entretanto el trabajo cognitivo continuado provoca aceleración de la FC y aumento de la presión sanguínea. Es por ello que las medidas psicofisiológicas de FC y presión sanguínea sistólica (PSS) y diastólica (PSD) han sido utilizadas para estudiar la facilidad o dificultad con que se desarrollan los procesos mentales (Coles y Duncan-Johnson, 1975; Holroyd, Westbrook, Wolf y Badhorn, 1978; Fernández y Vila, 1989). Se ha encontrado una relación moduladora entre variabilidad de la actividad cardiovascular y actividad mental (Mudler y Mudler, 1981; Wölk y Velden, 1987), así como un aumento de la FC y de la PSS ante distintas situaciones de laboratorio (Grizb, Quirós y Briales, 1993) y un aumento mayor de la FC durante la realización de una tarea cognitiva que antes o después de la misma (Harrell y Clark, 1985).

Nuestro propósito fue estudiar la posible existencia de un patrón de reacción cardiovascular específico del esfuerzo cognitivo. Para ello, registramos la actividad cardiovascular de una muestra de estudiantes universitarios en distintas situaciones de laboratorio (entre las que se incluía una tarea de evaluación) y naturales (entre las que se incluía una situación real de examen). Esperábamos encontrar mayor activación cardiovascular en los sujetos antes y durante la realización de un examen en comparación con un grupo de control. No esperábamos encontrar diferencias en activación cardiovas-

* Este trabajo es parte de una investigación más amplia subvencionada parcialmente por una Beca de Formación del Personal Investigador de la Consejería de Cultura, Educación y Turismo de la Comunidad Autónoma de la Región Murcia y por una ayuda del Vicerrectorado de Investigación de la Universidad de Murcia (Fondos FEDER, 1991).

cular entre las pruebas de evaluación en laboratorio y en la situación natural de examen, pero sí mayor activación cardiovascular en laboratorio durante la condición de evaluación en comparación con el resto de condiciones.

MÉTODO

Sujetos

56 estudiantes voluntarios, de las licenciaturas de Pedagogía y Psicología de la Universidad de Murcia, fueron seleccionados aleatoriamente entre todos los sujetos no ansiosos (según puntuaciones en el Cuestionario de Ansiedad Estado-Rasgo de Spielberger, Gorsuch y Lushene, 1986) a partir de un sondeo previo realizado a un total de 636 alumnos. Dicha muestra fue dividida en dos grupos de 28 sujetos (14 mujeres y 14 hombres): grupo «examen» (experimental), formado por alumnos que fueron citados para el experimento una hora antes de realizar un examen correspondiente a una asignatura de su licenciatura; y grupo «no-examen» (control), formado por alumnos que no tenían examen durante la semana en que realizaron el experimento.

Instrumentos

La FC y las PSS y PSD fueron registradas de forma automática mediante un monitor ambulatorio TM 2420/2020 (A&D Engineering, Milpitas, CA), homologado según criterios de la British Hypertension Society (White et al., 1991). Para grabar y reproducir las instrucciones a los sujetos, se utilizó un magnetófono TEAC-A-3440.

Procedimiento

Realizamos tres sesiones de registro de la actividad cardiovascular: dos en situación de laboratorio, iguales para los dos grupos de alumnos; y una en situación natural, intercalada entre las dos anteriores, en la cual los alumnos del grupo examen realizaban una prueba escrita eliminatória junto al resto de sus compañeros de curso, mientras registrábamos su actividad cardiovascular en cuatro ocasiones (una cada 15 minutos) durante la primera hora. La actividad cardiovascular de los alumnos del grupo no-examen fue registrada en intervalos temporales idénticos, mientras desarrollaban su actividad cotidiana.

Las dos sesiones de laboratorio (llevadas a cabo inmediatamente antes y después de la situación natural) se realizaron en el laboratorio de Psicofisiología Humana de la Universidad de Murcia, hallándose los sujetos sentados dentro de una cabina experimental semi-insonorizada y con la luz atenuada. Tras un período de adaptación de 3 minutos, comenzaba la sesión que constaba de cuatro fases o condiciones: una línea base (3 min.); seguida de las instrucciones en las que se pedía al sujeto que escuchase una lista de 20 nombres de 3 sílabas (10 propios y 10 comunes), ya que después se le preguntaría acerca de ellos; a continuación el sujeto permanecía en silencio durante tres minutos (fase de preparación) antes de recibir las instrucciones para escribir, durante dos minutos, los nombres que recordase (fase de evaluación); la sesión finalizaba con un descanso de tres minutos. La medida de la actividad cardiovascular se realizaba en la mitad de cada fase. La única diferencia entre las dos situaciones de laboratorio fue la composición de la lista de 20 nombres, aunque ambas eran semejantes en carga semántica e idénticas en longitud.

RESULTADOS Y CONCLUSIONES

Situación de Laboratorio

Los MANOVAs mixtos Grupo (examen vs. no-examen) x Sesión (primera vs. segunda sesión de laboratorio) x Condición (línea base, preparación, evaluación y descanso), realizados para cada una de nuestras variables cardiovasculares no muestran diferencias significativas para PSD, pero sí para PSS y FC. Se observa mayor PSS ($F_{1,54} = 6.86, p=.011$) y mayor FC ($F_{1,54} = 12.55, p=.001$) en ambos grupos durante la primera sesión de laboratorio, lo que puede interpretarse como un efecto de adaptación (ver Tabla 1). Como esperábamos, la PSS y la FC aumentan más durante la condición de evaluación que en las restantes ($F_{3,162} = 18.80, p<.001$, y $F_{3,162} = 57.15, p<.001$, respectivamente). Este resultado puede interpretarse como el efecto del esfuerzo cognitivo (Fernández y Vila, 1989; Grzib et al., 1993) y motor (Obrist, Howard, Lawler et al., 1974) requerido al sujeto.

TABLE 1
 MEDIAS (DESVIACIONES TÍPICAS) DE LA FRECUENCIA CARDÍACA (FC) Y LAS PRESIONES SANGUÍNEAS SISTÓLICA (PSS) Y DIASTÓLICA (PSD) PARA CADA GRUPO

	ANTES				DURANTE				DESPUÉS			
	L.Base	Prepar.	Ejecuci.	Descan.	L.Base	Prepar.	Ejecuci.	Descan.	L.Base	Prepar.	Ejecuci.	Descan.
<i>FC</i>												
<i>Examen</i>	74.46 (10.64)	74.82 (12.03)	89.93 (16.04)	76.07 (10.98)	91.90 (15.53)	70.93 (11.64)	72.25 (10.06)	78.54 (12.17)	71.82 (10.25)			
<i>No examen</i>	72.68 (13.98)	72.54 (12.75)	82.14 (16.82)	72.32 (12.33)	77.80 (13.63)	70.61 (13.44)	71.36 (14.18)	75.43 (13.15)	71.29 (13.12)			
<i>PSS</i>												
<i>Examen</i>	18.86 (12.23)	118.36 (11.32)	122.50 (15.09)	118.29 (12.46)	127.51 (14.47)	114.00 (11.54)	112.75 (10.13)	116.46 (10.59)	112.61 (10.76)			
<i>No examen</i>	12.32 (16.09)	112.82 (13.98)	117.86 (16.09)	110.29 (16.49)	121.21 (15.65)	110.82 (14.40)	110.61 (15.12)	115.14 (15.32)	110.50 (14.55)			
<i>PSD</i>												
<i>Examen</i>	75.61 (9.68)	77.86 (10.17)	77.96 (8.18)	76.50 (7.55)	79.05 (10.13)	75.43 (5.98)	75.71 (7.53)	77.54 (6.06)	76.50 (6.88)			
<i>No examen</i>	73.50 (9.98)	74.36 (9.74)	74.27 (10.53)	72.64 (9.46)	79.29 (13.66)	74.68 (10.40)	76.04 (10.24)	74.18 (9.29)	74.14 (9.32)			

Las diferencias entre nuestros dos grupos de alumnos sólo aparecen en PSS cuando analizamos por separado el comportamiento de cada uno de los grupos. El efecto Sesión sólo resultó significativo en el grupo examen ($F_{1,27} = 18.27$, $p < .001$), debido a una mayor PSS en la primera sesión de laboratorio (media= 119.5 mmHg; $dt = 12.77$) que en la segunda (media= 113.95 mmHg; $dt = 10.76$). Dado que este resultado no es significativo para el grupo de control, podemos atribuirlo, además de al efecto de adaptación, a una activación cardiovascular anticipatoria debida al examen que estos alumnos realizarían momentos después de nuestra primera sesión de laboratorio. Este resultado confirma parcialmente, al menos para la PSS, nuestra predicción de mayor activación cardiovascular antes del examen.

Situación Natural

La comparación de nuestros dos grupos en situación natural, mediante un ANOVA simple, sólo mostró diferencias significativas para la FC ($F_{1,54} = 13.03$, $p = .001$), siendo ésta mayor en el grupo examen (media= 91.90 lat/min, $dt = 15.53$) que en el grupo no-examen (media= 77.80 lat/min, $dt = 13.63$). Este resultado confirma, sólo para la FC, la hipótesis de Lacey (1967). El no encontrar las diferencias esperadas entre los grupos en presión sanguínea viene a decir que sólo la FC refleja una mayor activación en la situación de examen que en la situación natural de no-examen.

Comparación de las Situaciones de Laboratorio con la Situación Natural

Un MANOVA Grupo x Sesión (1ª sesión laboratorio x sesión natural x 2ª sesión de laboratorio) para cada una de las condiciones de registro en laboratorio (línea base, preparación, evaluación y descanso) por separado y para cada variable cardiovascular nos permitió comparar los niveles de actividad antes, durante y después de la situación natural. Resumiendo los resultados obtenidos, la actividad cardiovascular siempre fue mayor durante la situación natural que durante las situaciones de laboratorio (ver Tabla 1), salvo la PSD que mostró valores semejantes en situación natural y en las condiciones de preparación y evaluación en laboratorio. Por último, sólo en FC fue significativa la interacción Grupo x Sesión. Como esperábamos, el grupo examen mostró mayor FC antes y durante la situación natural (examen) que el grupo de control (no examen), mientras que en laboratorio no hay diferencias entre los grupos después de la situación natural, confirmándose parcialmente los resultados de Harrell y Clark (1985).

Conclusiones

Nuestros resultados permiten afirmar que el esfuerzo cognitivo durante pruebas de evaluación, tanto en el aula como en el laboratorio, se asocia a un aumento de la actividad cardiovascular. Como sabemos, estos valores cardiovasculares dificultan la distractibilidad ante estímulos externos a la tarea (Lacey, 1967). Futuras investigaciones deberán estudiar si ante pruebas de evaluación existen patrones diferentes de reactividad cardiovascular asociados a determinados tipos de deterioros en el rendimiento.

REFERENCIAS BIBLIOGRÁFICAS

- COLES, M. G. H. y DUNCAN-JOHNSON, C. C. (1975): Cardiac activity and information processing: The effects of stimulus significance, detection, and response requirements. *Journal of Experimental Psychology: Human Perception and Performance*, 1, 418-428.
- FERNÁNDEZ, M. C. y VILA, J. (1989): Cognitive versus motivational significance of the cardiac response to intense auditory stimulation. *International Journal of Psychophysiology*, 8, 49-59.

- GRIZB, G., QUIRÓS, P. y BRIALES, C. (1993): La tensión arterial como variable dependiente en psicología. *Revista de Psicología General y Aplicada*, 46, 161-169.
- HARRELL, J. P. y CLARK, V. R. (1985): Cardiac responses to psychological tasks: impedance cardiographic studies. *Biological Psychology*, 20, 261-283.
- HOLDROYD, K. A., WESTBROOK, T., WOLF, M. y BADHORN, E. (1978): Performance, cognition and physiological responding in test anxiety. *Journal of Abnormal Psychology*, 87, 442-451.
- KERRES, M. (1985): Objective and subjective arousal in test anxiety: differential accuracy of internal perception. En H. M. van der Ploeg, R. Schwarzer y Ch. D. Spielberger (Eds.): *Advances in test anxiety research*, vol. 4 (pp. 35-42). Lisse: Swets and Zeitlinger, B.V.
- LACEY, J. I. (1967): Somatic Response patterning and stress: Some revisions of activation theory. En M. H. Appley y R. Trumbull (Eds.): *Psychological stress: Issues in research* (pp. 14-42). Nueva York: Appleton Century Crofts.
- LACEY, J. I. y LACEY, B. C. (1970): Some autonomic-central nervous system interrelationships. En P. Black (Ed.): *Physiological correlates of emotion* (pp. 205-227). Nueva York: Academic Press.
- OBRIST, P. A., HOWARD, J. L., LAWLER, J. E., GALOSY, R., MEYERS, K. y GAEBELEIN, C. J. (1974): Cardiac-somatic interaction. En P. A. Obrist, A. H. Black, J. Brener y L. V. DiCara (Eds.): *Cardiovascular psychophysiology: Currents issues in response mechanisms, biofeedback and methodology* (pp. 136-162). Chicago: Aldine.
- MUDLER, G. y MUDLER, L. J. M. (1981): Information processing and cardiovascular control. *Psychophysiology*, 18, 392-402.
- REDONDO, M. y VALLE-INCLÁN DEL, F. (1992): Decrements in heart rate variability during memory search. *International Journal of Psychophysiology*, 13, 29-35.
- SPIELBERGER, C. D., GORSUCH, R. L. y LUSHENE, R. E. (1986): *Cuestionario de ansiedad estado-rasgo*. Madrid: TEA Ediciones S.A. (Edición original en inglés, 1970).
- WHITE, W. B., PICKERING, T. G., MONGANROTH, J., JAMES, G. D., MCCABE, E. J., MOUCHA, O. y HUNTER, H. (1991): A multicenter evaluation of the A&D TM-2420 ambulatory blood pressure recorder. *American Journal of Hypertension*, 4, 890-896.
- WÖLK, CH. y VELDEN, M. (1987): Detection variability within the cardiac cycle: toward a revision of the 'baroreceptor hypothesis'. *Journal of Psychophysiology*, 1, 61-65.

SER DE CIENCIAS..., SER DE LETRAS

Tomás Escudero Escorza
I.C.E. — Universidad de Zaragoza

Durante años, nuestros alumnos de los Cursos de Aptitud Pedagógica han obtenido información de calificaciones de evaluaciones escolares en distintas asignaturas, con vistas a realizar un trabajo práctico sobre correlación, en la asignatura de estadística para profesores. Por nuestra parte, hemos venido recogiendo las correlaciones presentadas, siempre que hubieran sido calculados a partir de un número suficiente de casos y que existieran garantías de que los datos fueran fiables.

A partir de parte de estas correlaciones, utilizando un sencillo análisis de la distribución de las mismas, pretendemos ofrecer alguna nueva información al debate-tópico al que hace referencia nuestro título y que, sin duda, entronca con un problema de investigación importante en el ámbito de la educación diferencial y la orientación. Nos estamos refiriendo lógicamente a la posible estructura factorial del rendimiento académico, alrededor de la que existe una cierta tradición investigadora en nuestro entorno educativo (González, 1988; Álvaro y otros, 1990; Vázquez, 1991; Vázquez y Manasero, 1992).

Hay que adelantar que este trabajo no incide en profundidad en ninguno de los ámbitos de fundamentación del tema, en él, simplemente, analizamos, presentamos y comentamos los resultados anteriormente aludidos.

1. NOTAS METODOLÓGICAS

Aunque el archivo de datos recogidos era mucho más amplio, en el análisis que presentamos nos hemos limitado a estudiar las correlaciones que incidieran de manera clara en el debate antes aludido. Por ello, nos hemos centrado en datos referidos a la E.G.B. y al B.U.P. y, dentro de ellos, a determinadas relaciones. En conjunto, hemos analizado 22 distribuciones distintas, con 710 casos (índices de correlación) en total.

Hemos de señalar que estos datos no están recogidos tras un diseño «ad hoc». Sin embargo, sí que responden a un proceso aleatorio. Durante años, cientos de alumnos de muy diversos lugares, han recogido correlaciones de multitud de clases, asignatura y cursos, en centros educativos de todo tipo, sin ningún tipo de orientación previa salvo en lo referente a la rigurosidad de la información.

Hemos partido de tres niveles educativos distintos E.G.B. (1º-5º cursos), E.G.B. (6º-8º cursos) y B.U.P.-C.O.U. Dentro de cada nivel, hemos estudiado las siguientes relaciones:

A. Rendimientos en matemáticas con rendimientos en: Ciencias, Lengua y Otras disciplinas.

B. Rendimientos en dos asignaturas: Ambas de Ciencias (esto solo en B.U.P.), ambas de Letras, una de Ciencias y otra de Letras.

C. Rendimientos *entre dos evaluaciones de la misma asignatura*: Ciencias o Matemáticas y Otras.

Con este esquema de ocho tipos de relaciones, se abarca de manera bastante completa todo el ámbito de las que inciden, o pueden incidir, en el análisis y explicación de la existencia o no de una tipología diferenciada de alumnos de ciencias y de alumnos de letras.

Si en el análisis de estas ocho distribuciones de datos encontramos tendencias sistemáticas diferentes para las relaciones, según la implicación de asignaturas de ciencias o de letras, nuestros datos apoyarán la hipótesis que plantea la existencia de una división en la tipología de los alumnos. Si no encontramos estas tendencias sistemáticas, nuestros datos apoyarán la hipótesis de que no existen diferencias en la estructura de rendimiento de los alumnos según una tipología de ciencias y otra de letras.

Este análisis de las distribuciones lo hemos llevado a cabo con el programa STATVIEW (Feldman y Gagnon, 1985). Hemos elaborado las tablas de frecuencias con intervalos de una décima de amplitud, representado los histogramas y calculando las estadísticas básicas para dicho análisis. Lógicamente y por razones de espacio, de toda esta información presentamos una síntesis de los aspectos que consideramos más relevantes.

2. RESULTADOS

En las tablas adjuntas se describen las 22 distribuciones agrupadas por cada tipo de correlación y dentro de cada tipo, los tres niveles educativos, con los datos ofrecidos, número de casos, media, desviación típica, rango, moda y asimetría, se tiene una idea bastante completa y precisa de la forma de cada distribución.

2.1. Correlaciones de los rendimientos en Matemáticas y en Ciencias

NIVEL	N	Me(r)	S _r	Máxim.	Mínim.	Moda (*)
E.G.B. (1-5)	3	0,64	0,23	0,80	0,30	
E.G.B. (6-8)	21	0,65	0,21	0,94	-0,01	0,60-0,70
B.U.P.-C.O.U.	70	0,66	0,21	0,97	-0,20	0,70-0,80

(*) Distribuciones con asimetría negativa

2.2 Correlaciones de los rendimientos en Matemáticas y en Lengua

NIVEL	N	Me(r)	S _r	Máxim.	Mínim.	Moda (*)
E.G.B. (1-5)	27	0,76	0,11	0,97	0,54	0,80-0,90
E.G.B. (6-8)	23	0,63	0,25	0,90	-0,18	0,80-0,90
B.U.P.-C.O.U.	33	0,39	0,23	0,89	-0,17	0,30-0,40

(*) Distribuciones con asimetría negativa en EGB, pero casi simétrica en B.U.P.-C.O.U.

2.3. Correlaciones de los rendimientos en Matemáticas y en Otras disciplinas (distintas de ciencias y lengua)

NIVEL	N	Me(r)	S _r	Máxim.	Mínim.	Moda (*)
E.G.B. (1-5)	6	0,61	0,26	0,91	0,16	0,70-0,80
E.G.B. (6-8)	32	0,64	0,24	0,97	-0,06	0,70-0,80
B.U.P.-C.O.U.	59	0,53	0,27	0,97	-0,10	0,70-0,80

(*) Distribuciones con asimetría negativa

2.4. Correlaciones entre los rendimientos en dos asignatura de Ciencias (sólo en BUP-COU)

NIVEL	N	Me(r)	S _r	Máxim.	Mínim.	Moda (*)
B.U.P.-C.O.U.	12	0,66	0,21	0,86	0,06	0,70-0,80

(*) Distribución con asimetría negativa

2.5. Correlaciones entre los rendimientos en dos asignaturas de Letras

NIVEL	N	Me(r)	S _r	Máxim.	Mínim.	Moda (*)
E.G.B. (1-5)	10	0,55	0,31	0,92	0,00	0,70-0,80
E.G.B. (6-8)	13	0,63	0,17	0,89	0,25	0,60-0,80
B.U.P.-C.O.U.	53	0,55	0,24	0,92	-0,18	0,60-0,70

(*) Distribuciones con asimetría negativa, algo menos acusada en B.U.P.-C.O.U.

2.6. Correlaciones entre rendimientos en una asignatura de Ciencias y Otra de letras

NIVEL	N	Me(r)	S _r	Máxim.	Mínim.	Moda (*)
E.G.B. (1-5)	8	0,62	0,39	0,24	-0,27	0,70-0,80
E.G.B. (6-8)	10	0,63	0,27	0,92	-0,02	0,70-0,80
B.U.P.-C.O.U.	37	0,40	0,25	0,84	-0,18	0,50-0,60

(*) Distribuciones con asimetría negativa, algo menos acentuada en B.U.P.-C.O.U.

2.7. Correlaciones entre dos evaluaciones de la misma asignatura (Ciencias y Matemáticas)

NIVEL	N	Me(r)	S _r	Máxim.	Mínim.	Moda (*)
E.G.B. (1-5)	7	0,73	0,27	0,94	0,14	0,80-0,90
E.G.B. (6-8)	24	0,72	0,13	0,92	0,40	0,70-0,80
B.U.P.-C.O.U.	132	0,66	0,17	0,95	0,08	0,70-0,80

(*) Distribuciones con asimetría negativa

2.8. Correlaciones entre dos evaluaciones de la misma asignatura (otras que Ciencias y Matemáticas)

NIVEL	N	Me(r)	S _r	Máxim.	Mínim.	Moda (*)
E.G.B. (1-5)	6	0,51	0,23	0,86	0,18	0,50-0,60
E.G.B. (6-8)	34	0,67	0,17	0,95	0,10	0,60-0,80
B.U.P.-C.O.U.	90	0,67	0,21	0,97	-0,26	0,70-0,80

(*) Distribuciones con asimetría negativa, más acusada en B.U.P.-C.O.U.

3. CONCLUSIONES Y DISCUSIÓN

Aunque no todas las distribuciones analizados recogen un número suficientemente amplio de casos, observamos que éstas responden mayoritariamente a una distribución con rango negativo y con un agrupamiento de casos alrededor de las correlaciones altas (0,60-0,80). Al mismo tiempo, existen casos aislados con correlaciones muy bajas e incluso negativas. Como es lógico, por el tipo de distribución, las medias son algo más bajas que las modas. Las desviaciones típicas ofrece cierta variabilidad, pero en muchos casos son bastante similares.

Las correlaciones del rendimiento en matemáticas con los rendimientos en ciencias, lengua y otras disciplinas ofrecen unas distribuciones muy parecidas en el nivel E.G.B. (6-8). La media de la correlación con lengua es algo más alta en E.G.B. (1-5) y algo más baja en B.U.P.-C.O.U.

Las correlaciones entre los rendimientos en dos asignaturas de ciencias, dos de letras y una de ciencias y otra de letras no ofrecen distribuciones muy diferentes, aunque se producen algunas fluctuaciones.

Las correlaciones entre rendimientos en dos evaluaciones de la misma asignatura no ofrecen ningún elemento que pudiera sorprendernos; son correlaciones altas y sin mucha variabilidad.

En general, los datos refuerzan la hipótesis de que hay un sustrato común en todos los rendimientos académicos. Si nos limitamos a los dos niveles de E.G.B., no encontramos ningún indicio que permita hablar de alguna posible diferencia entre ciencias y letras. Sin embargo, en el nivel B.U.P.-C.O.U., sí que se aparecen unas diferencias que parecen apoyar la hipótesis de distinción entre «alumnos de letras y de ciencias».

En B.U.P.-C.O.U., cuando hemos correlacionado los rendimientos de Ciencias y/o Matemáticas hemos obtenido sistemáticamente una distribución con media alrededor de 0,66, desviación típica alrededor de 0,20, moda alrededor de 0,75 y asimetría negativa.

Cuando hemos correlacionado dos rendimientos en otras asignaturas (ni ciencias, ni matemáticas) de B.U.P.-C.O.U., los datos son iguales cuando se trata de la misma asignatura, bajando un poco la media (0,55) cuando se trata de asignatura distintas.

Sin embargo, cuando hemos correlacionado rendimientos en ciencias y/o matemáticas con los de otras asignaturas de B.U.P.-C.O.U., las medias de la distribuciones han descendido de manera sensible en algunos casos (0,39; 0,53; 0,40), con ciertos cambios en la forma de la distribución.

En conjunto, nuestros datos parecen reforzar la hipótesis de una cierta diferenciación, aunque no muy acusada, entre alumnos de ciencia y de letras, pero solamente a partir de un determinado nivel educativo. Hasta el final de la educación obligatoria, no nos aparece más que un sustrato común para todo tipo de rendimientos académicos.

REFERENCIAS

- ÁLVARO, M. y otros (1990): *Hacia un modelo causal del rendimiento académico*. Madrid: M.E.C.—C.I.D.E.
- FELDMAN, D. y GAGNON, J. (1985): *Statview. The Graph Statistics Utility for the Macintosh*. Calabases, CA: BrainPower, Inc.
- GONZÁLEZ, A. J. (1988): Indicadores del rendimiento escolar. Relación entre pruebas objetivas y calificaciones. *Revista de Educación*, 287, 35-54.
- VÁZQUEZ, A. (1991): Estructura factorial del rendimiento académico en Bachillerato y C.O.U. *Bordon*, 43(1), 7-19.
- VÁZQUEZ, A y MANASSERO, M. A. (1992): Inteligencia y aptitudes en la predicción del rendimiento académico en matemáticas de Bachillerato. *Revista Española de Pedagogía*, 191, 153-180.

REVISTA
INVESTIGACIÓN
EDUCATIVA

BOLETÍN SUSCRIPCIÓN

Para suscribirse llene este boletín y devuélvalo a:

REVISTA DE INVESTIGACIÓN EDUCATIVA

Dpto. «Métodos de Investigación y Diagnóstico en Educación»

Facultad de C. Educación C/ Baldiri i Reixach, s/n, Bloque D, 3.º
08028 - BARCELONA

Nombre

D.N.I. o N.I.F.

Dirección

Población C.P.

País Teléfono

Coste de la inscripción:

Individual: 3.500 ptas.

Institucional: 5.000 ptas.

Números sueltos: 2.000 ptas.

Indicar n.º deseado:

Números extras: 2.500 ptas.

Indicar n.º deseado:

(Fecha y Firma)

BOLETÍN DE DOMICILIACIÓN BANCARIA

Señores,

Les agradeceré que con cargo a mi cuenta/libreta atiendan los recibos que les presentará la *Revista de Investigación Educativa*, como pago de mi suscripción a la misma.

Titular de la cuenta

Banco/Caja

N.º de cuenta N.º de libreta

Agencia

Población

(Fecha y Firma)

Para asociarse llene las dos partes de este boletín y devuélvalo a:

A.I.D.I.P.E.

Dpto. «Métodos de Investigación y Diagnóstico en Educación»

Facultad de C. Educación

C/ Baldiri i Reixach, s/n, Bloque D, 3.º

08028 - BARCELONA

Cuota de suscripción anual 5.000 ptas.

DATOS PERSONALES

Nombre y Apellidos

D.N.I. o N.I.F.

Dirección

Población C.P.

Provincia Teléfono ()

Deseo asociarme desde el día de de 19.....

DPTO. TRABAJO **CENTRO TRABAJO**

Situación profesional Dist. Universitario

DATOS BANCARIOS

Titular de la cuenta

Banco/Caja

Domicilio Agencia

Población C.P.

(Firma)

CODIGO CUENTA CLIENTE			
Entidad	Oficina	D.C.	Núm. de Cuenta

Señores,

Les agradeceré que con cargo a mi cuenta/libreta atiendan los recibos que les presentará la **Asociación Interuniversitaria de Investigación Pedagógica Experimental** como pago de mi cuota de asociado.

Titular de la cuenta

Banco/Caja

Domicilio Agencia

Población C.P.

(Fecha y Firma)

CODIGO CUENTA CLIENTE			
Entidad	Oficina	D.C.	Núm. de Cuenta