

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

ASOCIACIÓN INTERUNIVERSITARIA
DE INVESTIGACIÓN PEDAGÓGICA
EXPERIMENTAL

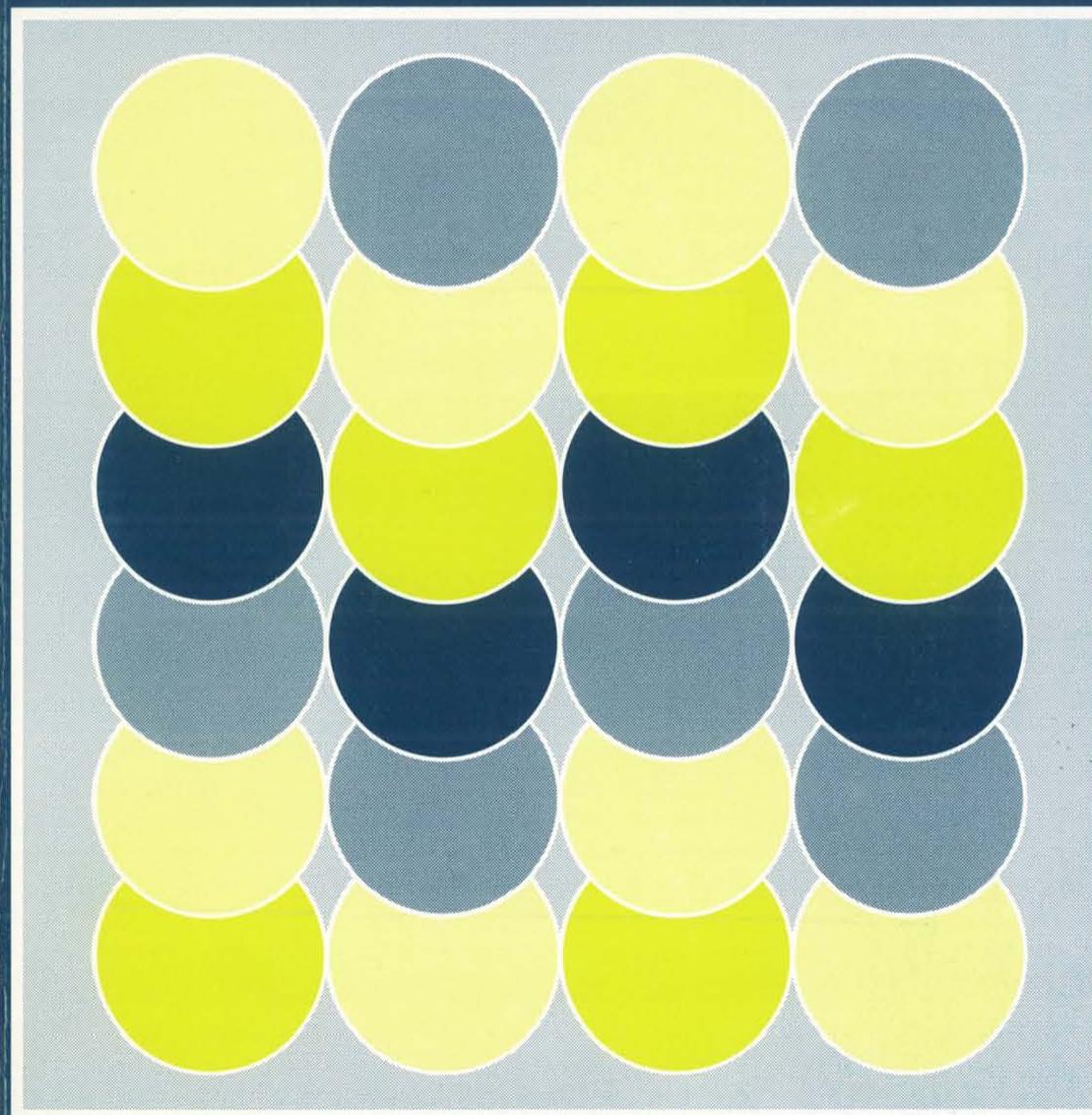
REVISTA DE INVESTIGACIÓN EDUCATIVA

N.º 21, 1.º semestre 1993



REVISTA
INVESTIGACIÓN
EDUCATIVA

N.º 21, 1.º semestre 1993



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 25 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

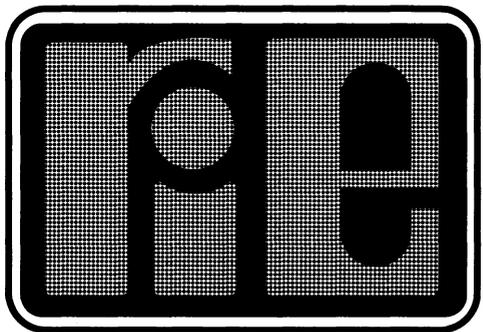
AIDIPE organiza cada 2 años, un Seminario Nacional con un tema central en torno al cual giran las ponencias, comunicaciones, mesas redondas. Los primeros Congresos han sido los de Barcelona (1981), Sitges (1983), Gijón (1985), Santiago de Compostela (1988), Murcia (1990).

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



Número: 21
1.º semestre, 1993

EDITA

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

Director: **Arturo de la Orden**

Directora ejecutiva:

Flor Cabrera

CONSEJO ASESOR:

Margarita Bartolomé

Nuria Borrell

Leonor Buendía

Iñaki Deandaluce

Lisardo Doval

Narciso García

Fuensanta Hernández

Jesús Jorner

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

Trinidad Donoso

Benito Echeverría

Julia V. Espín

Pilar Figueras

Ángel Forner

Javier Gil

Fuensanta Hernández

M.ª Luisa Rodríguez

Mercedes Rodríguez

Delio del Rincón

Antonio Sans

DISTRIBUCIÓN:

Área M.I.D.E.

Facultad de Educación

Campus Espinardo

Universidad de Murcia 30007

Tels. (968) 83 30 00 - 83 10 00

Ext. 2633

SUSCRIPCIÓN E

INTERCAMBIO CIENTÍFICO:

Dpto. MIDE

Facultad de Educación

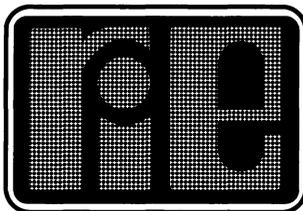
Baldiri Reixach, s/n. Bld. D-3.º

08028 BARCELONA

DISEÑO Y MAQUETACIÓN:

Compobell, S.L.

MURCIA



N.º 21

REVISTA INVESTIGACIÓN EDUCATIVA

SUMARIO

EDITORIAL	5
TRABAJOS DE INVESTIGACIÓN	
Modelo de Evaluación externa de la eficacia de los centros educativos por Blanca Rodríguez Díez	7
Análisis y valoración de programas pertenecientes al plan experimental para la reforma de la educación infantil por Mª Teresa Aguado Odina	29
Modelo para una autoevaluación modular en una institución educativa (B.A.D.I.) por Luis Blanco Felip	47
La posición del profesorado ante el cambio educativo. Un escalamiento multidimensional no métrico de los discursos sobre la reforma por Javier Gil Flores	67
ESTUDIO MONOGRÁFICO	
La competencia social como exigencia escolar: técnicas y procedimientos de evaluación por Gregorio Rodríguez, Eduardo Gª Jiménez y Carmen Gª Pastor	83
TRABAJOS METODOLÓGICOS	
Modelos borrosos en la medición escolar por Xavier Gil Quesada y Joan Mateo Andrés	107
Estudio de las calificaciones escolares mediante análisis exploratorio de datos por Lluís Salafranca Cosials, Montserrat Freixa Blanxart y Fco. Javier Ormazábal Unzué	125
Problemas fundamentales del análisis logarítmico lineal (y II): casillas vacías y casillas extremas por Ana Delia Correa Piñero	137
FICHAS-RESUMEN	151

EDITORIAL

De nuevo, durante este año, los socios de IDIPE nos encontraremos con motivo del VI Seminario sobre Modelos de Investigación Educativa. Fue en 1981 cuando nos encontramos por primera vez movidos por la necesidad de poner en común e intercambiar nuestras experiencias y hallazgos en el ámbito de la investigación educativa. En el transcurso de estos años hemos ido cumpliendo nuestras citas: recordemos nuestros primeros momentos en Sitges en el año 1983; en Gijón donde trabajamos *Nuevas tendencias de la investigación sobre Rendimiento Académico* en el año 1985; posteriormente en Santiago con el tema de *Evaluación de Centros* en el año 1987; y la última de ellas en Murcia, en 1990, reflexionando sobre *Metodologías en el Diagnóstico y Evaluación en los procesos de intervención educativa*.

El próximo seminario, bajo el tema general de *Investigación sobre la diferenciación educativa y orientación: la dimensión metodológica*, intenta responder a la necesidad constatada en la comunidad pedagógica universitaria de precisar e impulsar la metodología de investigación en el sector de la orientación educativa y en general en la intervención pedagógica diferenciada. Está abierto a los miembros de AIDIPE dedicados al estudio e investigación de problemas educativos en la perspectiva experimental y diferencial y a los profesionales del diagnóstico, la orientación y a diferentes niveles y modalidades de diferenciación educativa, así como a cuantos estén interesados en el desarrollo de la educación en esta línea.

El seminario se ha estructurado en cinco Ponencias que constituyen el núcleo temático en torno al cual podrán presentarse y discutirse las Comunicaciones de los asistentes.

Las Ponencias son las siguientes:

- I. Modelos de investigación en la intervención educativa diferencial.
- II. Nuevas aportaciones en la evaluación de programas de intervención psicopedagógica (Planes, estrategias, procedimientos, materiales, etc.).
- III. Modelos de construcción y validación de instrumentos diagnósticos.
- IV. Análisis de datos cualitativos en la investigación sobre la diferenciación educativa.
- V. Perspectivas metodológicas en la investigación sobre adaptación curricular e intervención psicopedagógica.

También se prevé la celebración de algunas Mesas Redondas con la participación de especialistas, se contempla la existencia de espacio y tiempo

suficientes para la presentación de pósters y la posibilidad de exponer libros o trabajos que hayan publicado los socios de AIDIPE en los últimos cinco años.

Os esperamos en Madrid los días 23, 24 y 25 de septiembre próximos. Será otra nueva ocasión para vernos y renovar nuestras ilusiones en torno a la asociación.

Nota Editorial: la revista agradece a los miembros del anterior Consejo Asesor la eficiente colaboración que ha encontrado en ellos, y notifica que los cambios realizados en el Consejo obedecen al interés por favorecer la sucesiva participación de los departamentos estableciendo una cierta rotación entre sus miembros.

TRABAJOS DE INVESTIGACIÓN

Revista Investigación Educativa - N.º 21 - 1993 (P. 7-27)

MODELO DE EVALUACIÓN EXTERNA DE LA EFICACIA DE LOS CENTROS EDUCATIVOS

por
Blanca Rodríguez Díez

RESUMEN

En este trabajo se aborda la evaluación de la eficacia de la escuela desde una perspectiva multivariable. Sigue la línea de las investigaciones que ponen de manifiesto que la eficacia no es el resultado de una serie de factores aislados, sino fruto de la combinación compleja de éstos. En el estudio se considera los siguientes factores: el modelo organizativo, los procesos que se desarrollan en el aula y el clima institucional.

ABSTRACT

This work contemplates the efficacy of school from a multivariate perspective. It is on the line of the searches devoted to put in that the efficacy it not the result of a serie of isolated factors, but the consequence of the complex combination of them. In the study, the following factors are considered: organizational model, proresses developed in school room and institutional climate.

La evaluación de la calidad de los procesos y los logros de los centros docentes como resultado de la acción educativa constituye una preocupación social en nuestros días.

La evaluación de la institución educativa puede ser considerada como el proceso de análisis científico y sistemático de las diversas variables de la escuela (el contexto, la estructura organizativa, el liderazgo del director, las prácticas de los profesores, la calidad del currículum, el clima, la innovación, la cultura, los recursos, la capacidad de resolver sus propios problemas, la calidad y cantidad de los logros

conseguidos, etc.), para realizar un diagnóstico científico, utilizando las herramientas de la investigación, que nos permitan formular juicios sobre su adecuación a unos criterios de calidad, previamente establecidos: el rendimiento satisfactorio de los alumnos, la calidad de vida de la institución, la satisfacción de alumnos, padres y profesores, etc.; en función de los cuales mejorar la toma de decisiones, diseñar un plan de intervención adecuado que ayude al centro a planificar y llevar a cabo estrategias de cambio y mejora de aquellos factores o áreas problemáticas. A la vez que permite dimensionar los logros conseguidos y consolidarlos.

Sin embargo, la evaluación de los centros docentes, realidad compleja y dinámica, constituye un proceso complejo, tanto por la multitud de variables susceptibles de evaluación como por la dificultad de evaluar variables procesuales que inciden simultáneamente en la eficacia de la institución escolar.

La complejidad de este proceso exigirá a los evaluadores elaborar y utilizar modelos, diseños y procedimientos de evaluación adecuadamente validados.

1. CARACTERÍSTICAS ORGANIZATIVAS DE LOS CENTROS EFICACES: MODELO ORGANIZATIVO, PROCESOS DE AULA Y CLIMA DE CENTRO

Los centros son los marcos donde tienen lugar la acción educativa. El análisis de la eficacia de los mismos exigirá, por una parte, el estudio de los «logros» alcanzados y, por otra, el estudio de cada uno de los factores que componen y actúan en la institución.

Desde la perspectiva de la investigación, la llamada «Effective school research», ha tratado de identificar las variables escolares relacionadas con la eficacia de los centros. Los factores de los mismos asociados con la eficacia, medida generalmente, a través del rendimiento de los alumnos, son diversos. Cada investigación sugiere alguno: metas y objetivos claros, factores de liderazgo, altas expectativas y estándares para los estudiantes, trabajo en equipo de profesores, etc. (Rutter, 1979; Purkey y Smith, 1982; Mackenzie, 1983; Brophy, 1981, 1986; Berliner, 1985), involucramiento de los padres, sistema de disciplina, sistemas de recompensas, desarrollo profesional del profesorado —staff development—, coordinación del currículum (Good y Brophy, 1986). Witte (1987) señala que la eficacia institucional depende de una combinación de factores de procesos que denomina «Clima Institucional».

Estas investigaciones han puesto de manifiesto que la eficacia no está determinada por la influencia aislada de cada uno de estos factores, sino por la combinación compleja de los mismos, dando lugar a unas características distintivas que hemos tratado de investigar en los centros educativos de nuestro contexto, ya que carecemos de investigaciones básicas sobre el funcionamiento real de los mismos. En nuestro estudio hemos considerado los siguientes factores: el modelo organizativo, los procesos que se desarrollan en las aulas y el clima del centro.

Modelo organizativo

Las estructuras organizativas —Órganos de gobierno y Equipos de profesores— en los centros eficaces deberían:

1. Posibilitar la participación colegiada y operativa de la Comunidad Educativa. La participación es uno de los elementos que más contribuye al desarrollo de un clima positivo abierto y con gran influencia en la eficacia. Esta participación debe estar planificada adecuadamente por el Equipo Directivo (Greembatt, 1983; Hall, 1982).

2. El Consejo Escolar servir de «catalizador» que regule las expectativas de la Comunidad Escolar, ya que las metas exigidas a los centros son diversas y en conflicto. Cada sector tiene sus propias exigencias e intereses y, por tanto, el Consejo Escolar debe hacer de «filtro» de las distintas y cambiantes demandas que los diversos sectores de la Comunidad propugnan de los centros.

3. El Equipo Directivo debe ejercer un liderazgo efectivo. La actuación del Equipo Directivo es clave, ya que es el responsable de clarificar las metas y objetivos del centro a los diferentes sectores de la Comunidad para conseguir compromiso y vinculación hacia Proyecto Educativo del centro (Mackenzie, 1983; Barnes, 1984; Glatter, 1990, etc.).

4. Los Equipos de Profesores —Departamentos/Seminarios, Equipos de Etapa, Ciclo, Nivel u otros grupos de profesores— tienen un papel clave en la articulación, organización y coordinación del Proyecto Curricular del centro. Este trabajo colaborativo posibilita el desarrollo profesional del profesorado, el mejor aprovechamiento de los recursos y la vinculación del profesorado a objetivos y proyectos comunes (Rutter, 1983; Wynne, 1980; Lieberman y Miller, 1984; Purkey y Smith, 1983; Deal y Kennedy, 1983; Raid, Hopkins y Holly, 1987, etc.).

Los procesos que tienen lugar en las aulas

El desarrollo de un currículum de calidad exigiría:

1. La planificación articulación del currículum. El profesor ha de planificar en equipo el currículum, es decir, ha de elaborar diseños de actuación que tengan en cuenta los objetivos, la estructura lógica de los contenidos, las capacidades de los alumnos, la dinámica del grupo, los sistemas de evaluación, etc. (Gimeno, 1981; Zabalza, 1987; Clark y Yinger, 1980; Shavelson y Stern, 1983; Pérez Gómez, 1984; Novak, 1985, etc.).

Esta planificación ha de aplicarla de forma flexible y contextualizada, adoptando soluciones alternativas cuando así lo aconsejen las diversas circunstancias del aula.

2. El plan de actividades diseñado para el grupo de alumnos debe ser diversificado. Para lo cual, el profesor programará actividades alternativas que permitan a

todos los alumnos alcanzar proporciones adecuadas de éxito, de acuerdo a sus capacidades (Brophy, 1983; Rosenshine, 1983; Marlieve y Filby, 1985).

3. La adecuación de las estrategias de enseñanza utilizadas a los objetivos, actividades y estilos de aprendizaje de los alumnos. El profesor ha de ser capaz de adaptar sus estrategias, según se trate de adquirir habilidades básicas o procesos cognitivos más complejos y a los alumnos con diferentes capacidades (Peterson, 1978; Rosenshine y Steven, 1985; Joyce y Weill, 1985; Doyle, 1983, 1986).

4. Las estrategias de aprendizaje que se promueven en los alumnos han de permitir a éstos la adquisición de aprendizajes significativos y duraderos. Como reiteradamente han afirmado diversos teóricos (Piaget e Inherder, 1969; Delval, 1983; Coll, 1985; Ausubel, 1968, etc.) el alumno no es un pasivo receptor de estímulos, el conocimiento no es una simple copia de la realidad, el alumno necesita construirlo. El desarrollo de sus capacidades requiere la participación del alumno, activando conocimientos y esquemas ya constuidos en etapas anteriores que, como consecuencia de su ejercicio, se modifican y perfeccionan.

5. El clima del aula debe ser positivo y abierto. Éste es promovido por las altas expectativas por parte del profesor sobre las posibilidades educativas de todos los alumnos, la existencia de normas claras y una atmósfera de convivencia adecuada en el aula (Doyle, 1986; Pelli y Corwell, 1982; Evertson, 1982, etc.).

Clima institucional

El Clima del centro positivo y abierto generado por:

1. La participación de la Comunidad Educativa en la vida del centro. La participación y el involucramiento de profesores y padres en la vida del centro ha sido analizada en numerosas investigaciones y se ha revelado como una dimensión importante del clima institucional (Brekenridge, 1979; Phi Delta Kappa, 1980, etc.). Esta vinculación no significa participación sin más, sino que supone consensos de metas, objetivos y valores educativos propugnados por el centro (Brookover y Rigsby, 1973).

2. La existencia de un sistema efectivo de participación de los estudiantes en la vida del centro (Cox, 1978; Duke y Perry, 1978; Urich Batchelder, 1979; Breckenridge, 1973). La participación de los alumnos en el currículum escolar y extraescolar está fuertemente relacionada con el clima y la satisfacción del alumnado (Eps-tein y MacPartland, 1979; Rutter, 1979).

3. La potenciación de una cultura colaborativa entre los profesores y el Equipo Directivo. Ello implica el estudio, análisis, reflexión, comunicación y revisión conjunta de los principios educativos que promueven y los procesos que están utilizando para su consecución (Mehaffy, 1989; Hopkins, 1988; Lieberman, 1988; Holly y Southworth, 1989; Escudero, 1990).

4. Las normas de convivencia claras, razonables, conocidas y aceptadas por

todos. Estas normas, aplicadas constructivamente, promueven en los alumnos y profesores sentimientos de responsabilidad y vinculación hacia el centro. Además de estas normas formales, existen en los centros un conjunto de normas no explícitas, que se manifiestan en el comportamiento de los diferentes grupos de la institución. El conocimiento y el consenso de las mismas con la Comunidad educativa promueve la existencia de relaciones positivas entre los miembros de la misma (Walberg, 1979; Wynne, 1980; Licatta, 1979; Curwin y Mendler, 1980, 1985).

5. La existencia en el centro de una cultura que promueva las altas expectativas de los profesores acerca las posibilidades educativas de los alumnos, pues, además de ser un poderoso factor capaz, no sólo de promover el rendimiento de éstos, sino también de generar un clima positivo a nivel de centro (Brophy y Good, 1974; Singer, 1978; Brookover, 1979, etc.).

6. La satisfacción profesional del profesorado, es decir, la percepción que ellos tienen de las ventajas e inconvenientes en el desarrollo de su profesión, se ha evidenciado como una variable consistentemente asociada con el clima institucional (Kalis, 1980; Sargent, 1979).

2. ESTUDIO EMPÍRICO

En nuestro modelo hemos considerado como unidad de análisis el centro docente en su totalidad, teniendo en cuenta los factores que, tanto las investigaciones anteriores como nuestro conocimiento del contexto escolar, han destacado como condicionadoras de la eficacia de los mismos (ver Centra y Potter, 1980; Stoel, 1987; Seashore, 1987; Mateo y Rodríguez Espinar, 1984; Tejedor, 1985; De Miguel, 1984, 1988; Aurora Fuentes, 1985, etc.).

El modelo se basa en un conjunto de hipótesis relacionadas, expuestas anteriormente y sintetizadas en la siguiente ecuación general:

«La *eficacia* institucional, (*R*), medida a través de una serie de indicadores (rendimiento satisfactorio de los alumnos, grado de satisfacción de alumnos y familias con el centro, bajo índice de absentismo escolar, baja proporción de alumnos desfasados edad/nivel, capacidad de centro de compensar los déficits de los alumnos ligados al medio familiar, así como el número de solicitudes de ingreso que recibe) depende directamente de la existencia de un adecuado *modelo organizativo* —Órganos de gobierno colegiados (Equipo directivo, (*OED*), Consejo Escolar, (*OCE*), y Claustro, (*OCL*) y funcionamiento en equipo de los profesores, (*OEP*)— que utilice y genere los *recursos* necesarios, (*D*), para llevar a cabo su proyecto educativo; potencie un *clima* de interacciones positivas, (*CI*), entre los diferentes grupos del centro, para optimizar la calidad de *la enseñanza* de los currícula que se desarrollan en las aulas, (*PA*), teniendo en cuenta las características de los *alumnos*, (*A*), las personales y profesionales de los *profesores*, (*P*), así como del *tipo de centro*, (*TC*), donde se desarrolla la acción educativa».

La ecuación general estructural correspondiente es:

$$R = \gamma_{111} \cdot TC + \gamma_{112} \cdot D + \gamma_{113} \cdot P + \gamma_{114} \cdot A + \beta_{115} \cdot OED + \beta_{112} \cdot OCE + \\ + \beta_{118} \cdot OEP + \beta_{119} \cdot CI + \beta_{110} \cdot PA + \psi_{111} \cdot \xi_{11}$$

Donde γ y β representan los parámetros que relacionan las variables exógenas con las endógenas y éstas entre sí, respectivamente.

3. METODOLOGÍA

3.1. Selección de la muestra

A partir del Nomenclátor de datos disponibles en el Servicio de Inspección se seleccionó una muestra de 31 centros docentes de una población de 350 que existen en la provincia.

Se tuvieron en cuenta los siguientes criterios: centros mayores de ocho unidades tanto públicos como privados, de diferentes «status socio-culturales» (bajo, medio bajo, medio, medio alto y alto) elegidos mediante muestro aleatorio, estratificado y proporcional.

La tabla 1, recoge el número grupos y elementos encuestados y/o evaluados:

Grupos y elementos	Número
Número de centros	31
Equipos Directivos	31
Consejos Escolares	31
Equipos de Profesores	150
Clima Institucional	400
Aulas Observadas	186
Alumnos encuestados	2.800
Familias encuestadas	2.000

Figura 1. *Tabla de tamaños muestrales.*

3.2. Medida de variables

Las variables se agruparon en exógenas o independientes y endógenas o dependientes, que, a su vez, aparecen en dos subgrupos: de proceso y de producto o resultados.

La medición de las variables se ha llevado a cabo utilizando varios procedimientos: el modelo organizativo del centro —Órganos de gobierno y funcionamiento en equipo de los profesores— a través de la observación del funcionamiento de la estructura organizativa correspondiente por el equipo evaluador, utilizando las esca-

las «valoración del equipo directivo, valoración del consejo escolar, valoración del claustro de profesores, departamentos, equipos de ciclo y nivel», entrevistas y análisis documental. El clima del centro mediante el cuestionario anónimo «Clima de centro» aplicado a una muestra de profesores elegidos al azar. Los procesos del aula a través de la observación del trabajo desarrollado por los profesores en las aulas, utilizando la escala «Evaluación de procesos de aula». El nivel de satisfacción de alumnos y familias con el centro mediante los cuestionarios «Satisfacción de alumnos/as y satisfacción de padres», respectivamente, aplicados a una muestra de los mismos. Las escalas y cuestionarios han sido validadas.

3.3. Técnica de análisis

Para la contrastación empírica del modelo hemos utilizado la técnica del «path analysis».

Estas técnicas metodológicas permiten estudiar el patrón de causación entre las variables, especificadas en el modelo, para determinar su plausibilidad. Es decir, con ellas se validan las estructuras causales que hayamos planteado en el modelo en función de una teoría previa.

Si el modelo es congruente con los datos, lo más que se puede afirmar es que apoya la teoría que lo ha generado. Podemos, sobre esta base, rechazar los modelos, corregirlos y mejorarlos (Pedhazur, 1982; Maruyama y Walberg, 1982; Cliff, 1983; Visauta, 1986; Bisquerra, 1989, etc.), pero no podemos interpretarlos como una prueba de la teoría, puesto que con los mismos datos correlacionales pueden ser congruentes varios modelos.

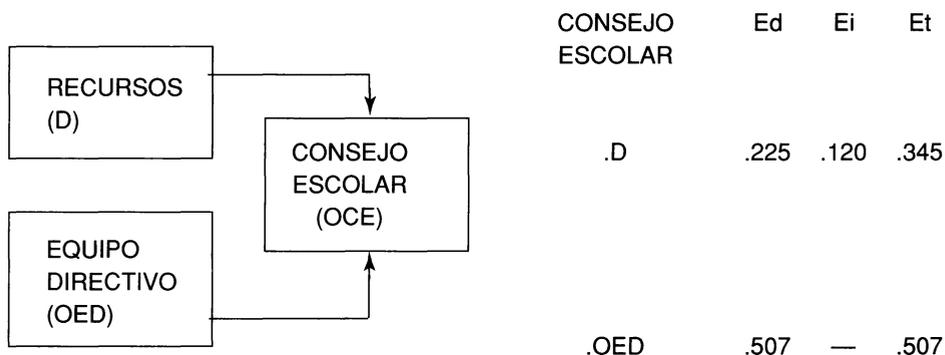
Los parámetros estructurales se calcularon mediante el programa informático LISREL VI (Jöreskog y Sörbom, 1985), utilizando el método de máxima verosimilitud, (ML), (ver Kerlinger y Pedhazur, 1973; Bisquerra, 1989, Byrne, 1989, etc.), que nos permitieron ir ajustando el primitivo modelo teórico hasta el definitivo.

4. RESULTADOS DEL ANÁLISIS CAUSAL

En la figura 2, (véase pág. 7) presentamos el path diagrama con el valor de los parámetros y los residuales obtenidos en el análisis. En los apartados siguientes analizaremos el efecto de las variables institucionales más importantes sobre otras, es decir, cada variable dependiente aparece con sus respectivas explicativas y las líneas continuas hacen referencia a una relación directa, mientras que las de trazos representan relaciones indirectas.

4.1. Tabla 2. Análisis causal: Valores de las variables sobre CONSEJO ESCOLAR, (OCE)

Diagrama de relaciones y valores de los efectos de sus variables explicativas:
 (—>Efecto directo) (—>Efecto indirecto)



El Consejo Escolar, (OCE), de acuerdo con el modelo propuesto, está influido por la variable Equipo Directivo, (OED), con un efecto directo $E_d = .507$, y por la variable recursos, (D), con los efectos $E_d = .225$, $E_i = .120$ y $E_t = .345$.

Los resultados obtenidos son congruentes con nuestras hipótesis ya que el funcionamiento del Consejo Escolar está influido en gran medida por el Equipo Directivo.

Hemos de tener en cuenta que el presidente de este órgano es el director del centro y los restantes componentes del Equipo Directivo (Jefe de Estudios y Secretario) son integrantes del mismo. El Equipo Directivo es quien reúne o convoca a dicha «unidad organizativa», determina los temas a tratar y establece, en la práctica, la dinámica de funcionamiento de la misma.

De hecho, resulta difícil que un Consejo Escolar, por la composición del mismo (mayor número de componentes profesionales que no profesionales), número de reuniones preceptivas en el curso, etc., pueda ejercer las atribuciones encomendadas por la L.O.D.E. en los centros escolares, si el Equipo Directivo no tiene voluntad de hacerlo funcionar.

Parece claro, pues, la influencia que el Equipo Directivo ejerce en el funcionamiento del Consejo Escolar de los centros.

De ahí parece evidenciarse que un Consejo Escolar fracasará en el ejercicio de sus atribuciones, si el centro no dispone de un Equipo Directivo profesionalmente eficaz.

La variable recursos, (D), tiene, sobre el Consejo escolar, un $E_t = .345$. Resulta congruente con nuestros planteamientos en base a que una de las atribuciones asignadas al Consejo Escolar, a la que en la actualidad éstos dan prioridad, es conseguir de la Administración Educativa recursos para los centros, circunstancia que explicaría esta relación.

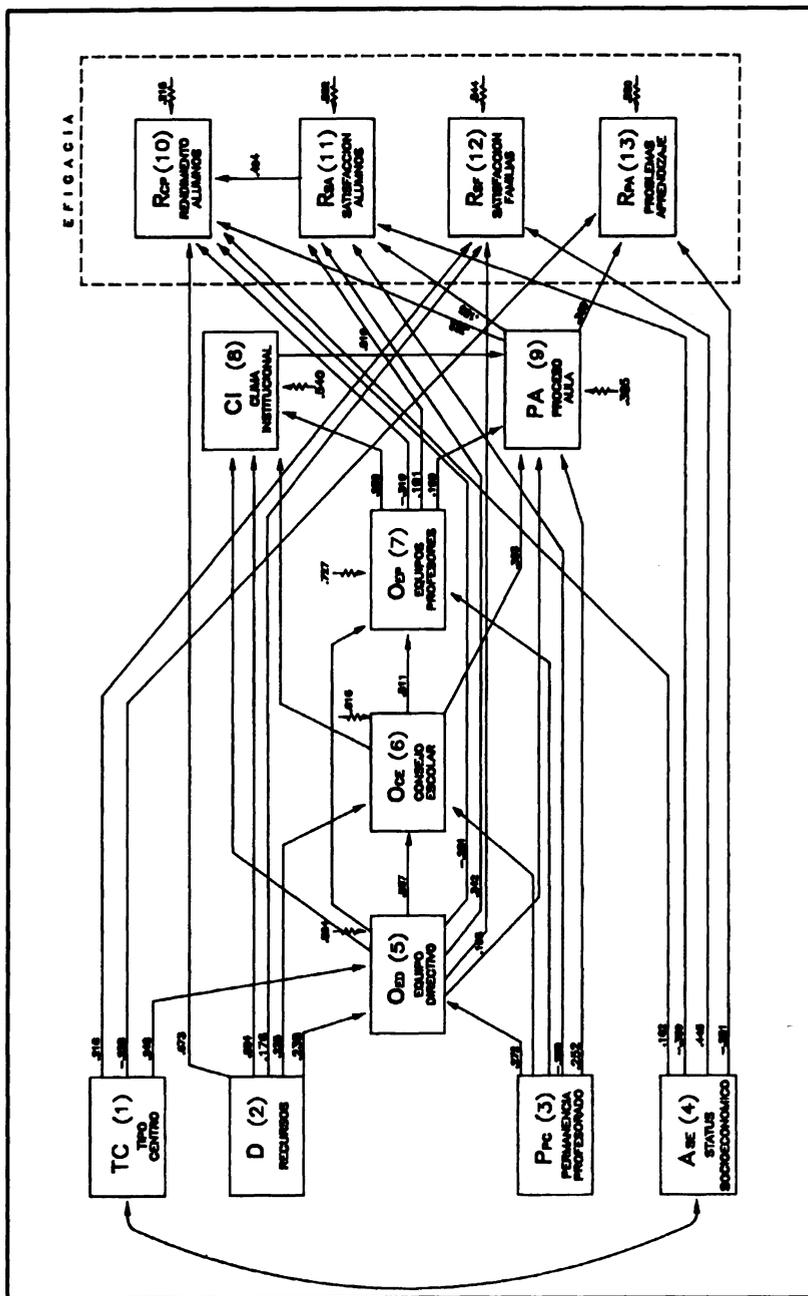
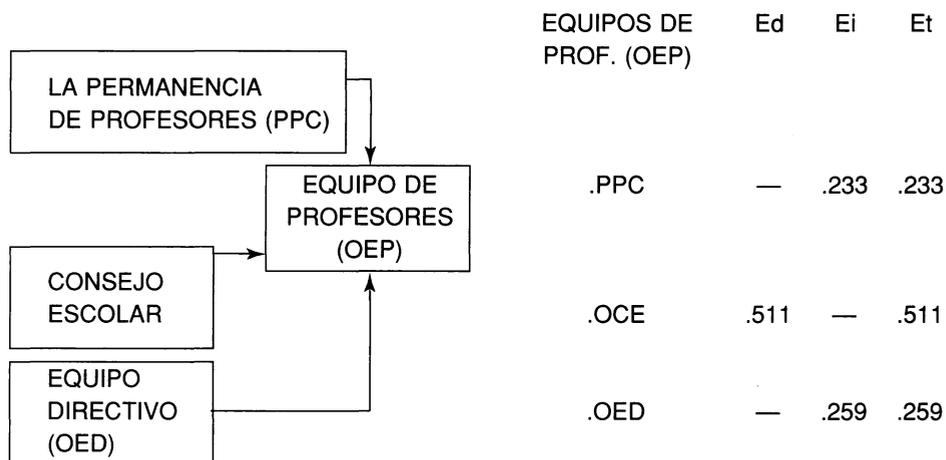


Figura 2. Path diagrama con los valores de los parámetros.

4.2. Tabla 3. Análisis causal: Valores de las variables sobre EQUIPOS DE PROFESORES (OEP)

Diagrama de relaciones y valores de los efectos de sus variables explicativas: (—>Efecto directo) (—>Efecto indirecto)



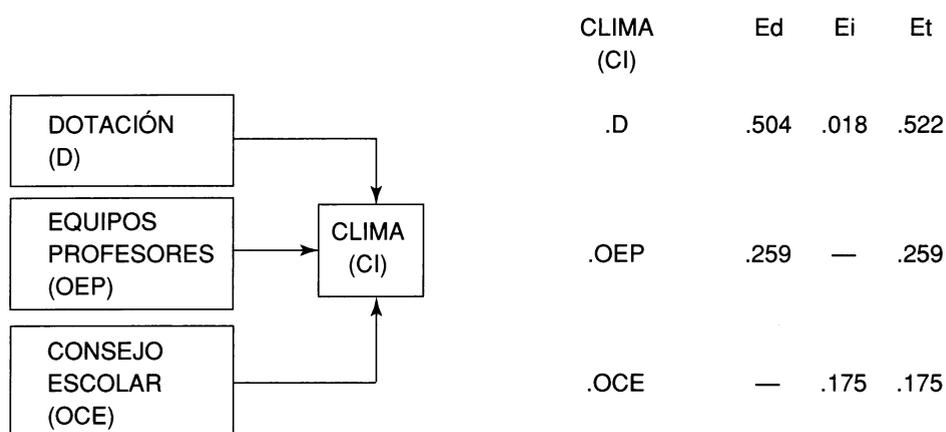
Los factores del centro que ejercen mayor influencia en orden a que los profesores trabajen en equipo son: en primer lugar, el Consejo Escolar, cuyo efecto directo es $Ed=.511$. En segundo lugar, el Equipo Directivo, (OED), con un efecto indirecto, a través de la variable Consejo Escolar, (OCE), de $Ei=.259$. En tercer lugar, la variable permanencia de los profesores en los centros, (PPC), con un efecto indirecto de $Ei=.233$.

El hecho de que sea el Consejo Escolar del centro el que ejerza mayor influencia, en lugar del Equipo Directivo, creemos se debe a que, en la práctica, ambas unidades organizativas intervienen sobre las mismas dimensiones de los centros. Así, el Equipo Directivo elabora el proyecto educativo del centro, pero es el Consejo Escolar quien lo aprueba y evalúa su consecución.

Esta capacidad del Consejo Escolar de valorar los logros del centro en diversos aspectos tales como: actividades docentes, rendimiento de los alumnos, currículum extraescolar, actividades económico-administrativas y resolución de problemas de convivencia, determina que tenga una influencia mayor, en algunas áreas del funcionamiento del centro, que el propio Equipo Directivo.

4.3. Tabla 4. Análisis causal: Valores de las variables sobre CLIMA DEL CENTRO (CI)

Diagrama de relaciones y valores de los efectos de sus variables explicativas: (—>Efecto directo) (—>Efecto indirecto)



Sobre esta variable habíamos hipotetizado influencia importante de las estructuras organizativas (Equipo Directivo, Consejo Escolar y Equipo de Profesores). Sin embargo, los resultados del análisis causal muestran que las variables de la escuela que ejercen mayor influencia en el clima general de la misma, (CI), son los llamados recursos escolares (instalaciones, equipamiento, materiales didácticos y la existencia en el centro de profesores especialistas), con un efecto directo, $Ed=.504$, y un $Et=.522$.

Hemos de considerar que, dentro de la variable recursos escolares, se ha incluido la existencia de materiales didácticos y de profesores de apoyo. El contar con especialistas que pueden ayudarles en sus funciones docentes, tiene, evidentemente, una alta importancia en el nivel de satisfacción y desarrollo de sus actividades.

El trabajo en equipo de los profesores, planificando y desarrollando el currículum, intercambiando experiencias y resolviendo conjuntamente los problemas comunes sobre la enseñanza, genera un clima positivo y un contexto enriquecedor para sus integrantes.

De ahí que, de las unidades organizativas, sea el funcionamiento de Equipos de Profesores la que aparezca con mayor valor de influencia $Ed=.259$.

De acuerdo con los resultados del análisis causal, también, la variable Equipo de Profesores es la «mediadora» de la influencia indirecta del Consejo Escolar sobre el clima, (CI), con un $Ei=.175$.

4.4. Tabla 5. Análisis causal: Valores de las variables sobre PROCESOS DE AULA (PA)

Diagrama de relaciones y valores de los efectos de sus variables explicativas: (—>Efecto directo) (—>Efecto indirecto).

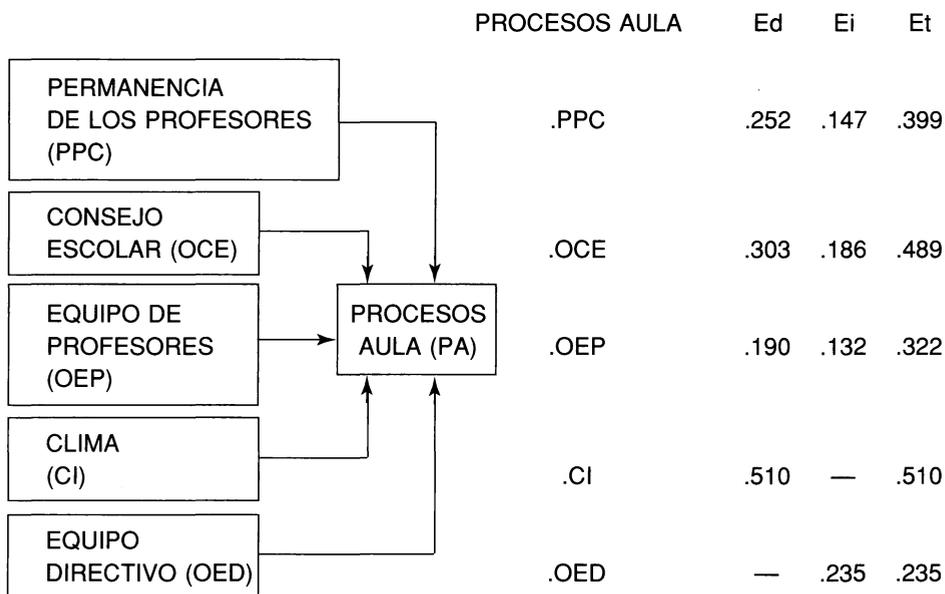
Las relaciones establecidas, a partir del análisis causal, nos evidencian que las variables a nivel de centro que ejercen una mayor influencia en los procesos de

enseñanza-aprendizaje que tienen lugar en las aulas son: el Clima global del centro, (CI), con un efecto directo $E_d=.510$; El Consejo Escolar, (OCE), que ejerce unos efectos $E_d=.303$, $E_i=.186$, $E_t=.489$; la Permanencia de los profesores en los centros, (PPC), dada por $E_d=.252$, $E_i=.147$, y $E_t=.399$; Equipo de Profesores, (OEP), cuyos efectos son $E_d=.190$, $E_i=.132$ y $E_t=.322$.

Por último, el Equipo Directivo, (OED), ejerce sobre los procesos que tienen lugar en las aulas, (PA), un efecto indirecto de $E_i=.235$, a través del Consejo Escolar.

De estas relaciones nos interesan destacar varios aspectos:

Los procesos que tienen lugar en las aulas (calidad del currículum impartido por los profesores, ayuda y orientación al alumnado, así como la gestión del aula y el clima que el profesor logre establecer entre él y sus alumnos y dentro del grupo), están influidos en gran medida por el clima general global que se establece en el centro. De ahí el alto valor, ($E_i=.510$), observado.



Estos resultados son coherentes con nuestro planteamiento teórico, ya que parece evidente el hecho de que un centro que logra generar un clima de trabajo positivo, producto de la participación de los profesores en la vida del centro y del trabajo cooperativo, tenga gran influencia en la calidad de los procesos que los profesores desarrollan en las aulas.

Este proceso causal evidencia que el marco de relaciones que se establecen a nivel general del centro promueve y configura las interacciones que se generan a nivel de aula. Es decir, el conjunto de relaciones asumidas por el centro son comunicadas y trasvasadas a las aulas.

De forma recíproca, las relaciones existentes dentro del aula, repercuten sobre el clima institucional general, produciéndose, probablemente, aunque no se haya planteado una recursividad entre variables, unos procesos complementarios e interdependientes.

Así mismo la mayor estabilidad de los profesores en los centros, (PPC), permite a los mismos, además de una mayor facilidad para trabajar en equipo, poder hacer planteamientos comunes sobre los procesos enseñanza-aprendizaje, así como revisar la validez de los mismos, por lo que un $E_t = .399$, parece lógico.

Los Equipos de Profesores, (OEP), ejercen, en buena parte, sus efectos indirectamente a través del Clima. Los datos del análisis nos parecen congruentes con el funcionamiento real de los centros, en los que los profesores trabajan formando equipos, más por afinidades personales que dentro de unas estructuras normativas establecidas.

Así, a partir del análisis del funcionamiento real de estos equipos se constata que, en general, aunque hayan elaborado un plan de trabajo formal (objetivos, actividades, calendario, responsable, etc.) el desarrollo real está muy alejado de las definiciones formales.

La influencia del Equipo Directivo sobre los procesos que tienen lugar en las aulas (todos indirectos) también parecen congruentes con el funcionamiento real, pues raramente interviene directamente en el trabajo de los profesores en las aulas, sino que, generalmente, lo hace a través de los Equipos de Profesores, de su contribución a la creación de un clima positivo y del Consejo Escolar, que aprueba el plan de centro, del que forman parte las programaciones curriculares y los criterios de evaluación y recuperación utilizados por los profesores en sus aulas.

Este proceso complejo que se establece entre las variables organizativas a nivel centro y su influencia en los procesos que tienen lugar en las aulas, esto es, que sea el OCE y no el OED quien ejerza una mayor influencia en PA, concuerda con el funcionamiento real y evidencia el «desacoplamiento» entre la unidad organizativa y la actividad a desarrollar por ella.

En consecuencia, las escuelas ofrecen la «imagen de anarquías organizadas», «débilmente articuladas», pues, aunque las unidades organizativas son interdependientes, esta interdependencia es relativa, ya que dichas unidades mantienen entre sí grandes niveles de autonomía funcional.

En este sentido, aunque en los centros existe una reglamentación que establece la coordinación entre los profesores (planes de acción de los departamentos, ciclos y niveles, etc.), estas reuniones son, con frecuencia, más para transmitir información y tratar temas burocráticos que verdaderas sesiones de trabajo sobre el currículum a desarrollar. Esta ausencia de interacción profesional resta a los profesores, en palabras de Sergiovanni (1987) oportunidades de buscar y recibir ayuda en los problemas que se les presentan en su actividad diaria y hace que la existencia de dichos equipos afecte mínimamente a las prácticas desarrolladas por los profesores en sus aulas.

4.6. Tabla 6. Análisis causal: Valores de las variables sobre SATISFACCIÓN DE LOS ALUMNOS (RSA)

Diagrama de relaciones y valores de los efectos de sus variables explicativas: (—>Efecto directo) (—>Efecto indirecto).

El Grado de satisfacción, (RSA), que manifiestan los alumnos sobre la calidad de las actividades curriculares y extracurriculares, la orientación y ayuda recibida por parte de los profesores, la actitud de los alumnos hacia la escuela, así como la calidad de las interacciones entre alumnos-profesores y entre los mismos alumnos, está condicionada por: el status socioeconómico, (ASE), el Consejo Escolar, (OCE), el Equipo de Profesores, (OEP), los procesos que tienen lugar en las aulas, (PA), y la permanencia del profesorado, (PPC).

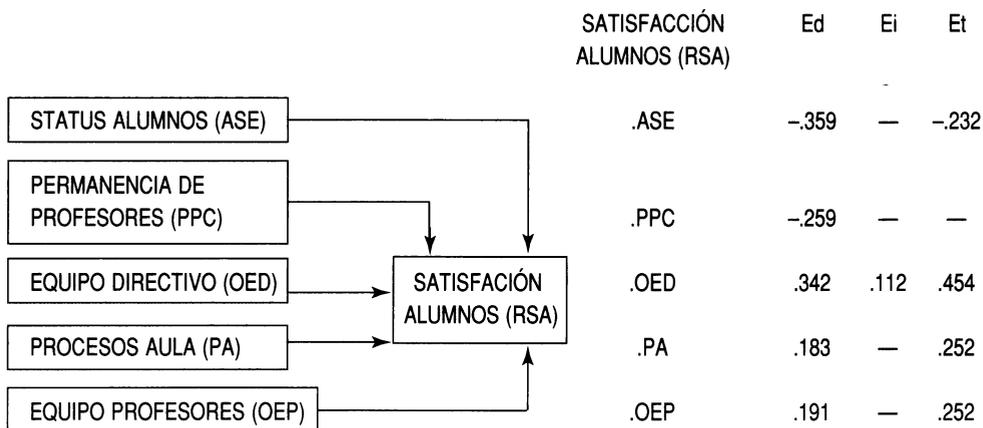
La clase social del alumnado, (ASE), ejerce un efecto directo y negativo sobre el grado de satisfacción del mismo, (RSA), de $Ed=.359$. La relación que se establece entre estas variables pone de manifiesto que los alumnos de status socioeconómicos más alto muestran menor satisfacción con su centro.

Estos resultados se explican, porque son los padres los que eligen, tanto más cuanto sus posibilidades se lo permiten, el tipo de enseñanza que desean para sus hijos.

Esta elección toma como indicador de calidad, en general, un alto nivel de exigencia de conocimientos a adquirir por sus hijos, énfasis hacia las actividades académicas, calidad de las actividades extracurriculares, ambientes ordenados y controlados por directores y profesores.

Estos resultados evidencian (a partir del análisis de los cuestionarios aplicados) que la insatisfacción de los alumnos de «status socioeconómico» alto con la escuela, se debe al alto nivel de exigencia del centro en el programa de actividades propuestas y a las «excesivas expectativas» que tienen los profesores sobre el rendimiento de los alumnos.

La variable, (OED), muestra unos valores altos sobre el grado de satisfacción de los alumnos, (RSA) $Ed=.342$, $Ei=.112$, $Et=.454$.



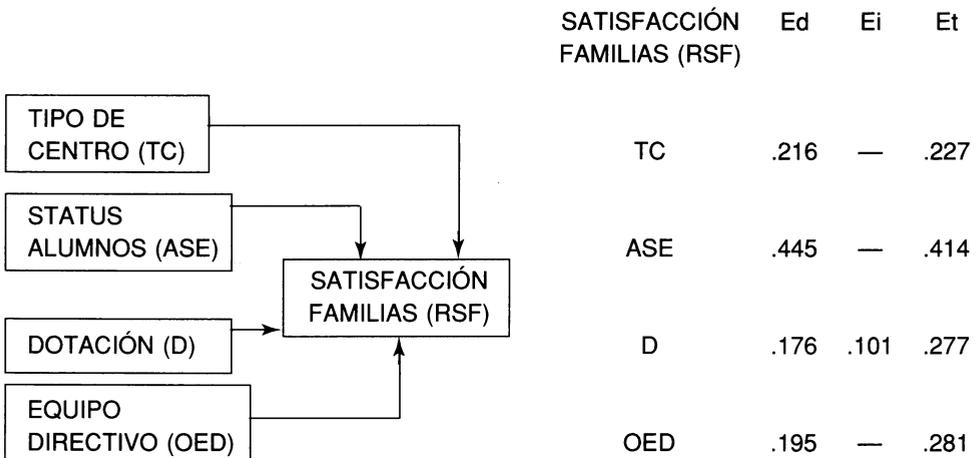
Los resultados son congruentes con nuestras hipótesis y confirman las particulares relaciones que esta variable mantiene con el resto de las estructuras organizativas, pues, si bien la relación directa del alumno se establece con sus profesores y, en principio, sería ésta la que debería manifestarse como la de mayor influencia, es el Equipo Directivo, quien tiene a su cargo las funciones de disciplina, de participación y de planificación y coordinación de las actividades extraescolares, de gran importancia para la satisfacción de los alumnos.

A los procesos que se desarrollan en las aulas, (PA), les habíamos supuesto un efecto mayor sobre la satisfacción de los alumnos; los resultados del análisis ponen de manifiesto que son bajos: $Ed=.183$.

Estos valores han de analizarse a la vista de los hallados para otra variable de profesorado, permanencia de los profesores en los centros, (PPC), cuya influencia sobre la satisfacción de los alumnos es: $Ed=.259$, significando que los profesores evaluados como más eficaces y con más experiencia, no producen una mayor satisfacción en los alumnos, en el sentido de que se muestran insatisfechos con la escuela consideran que el nivel de exigencia es elevado, nivel de exigencia que parece ser mayor cuando el profesor es más experimentado.

6.7. Tabla 7. Análisis causal: Valores de las variables sobre SATISFACCIÓN FAMILIAS (RSF)

Diagrama de relaciones y valores de los efectos de sus variables explicativas: (—>Efecto directo) (—>Efecto indirecto)



En nuestro modelo teórico habíamos supuesto que el nivel de satisfacción de las familias estaría condicionado por el tipo de centro, su status social, la acción del

Equipo Directivo y la calidad de los procesos que se desarrollan en las aulas. Las relaciones que se establecen a partir del modelo causal son:

El tipo de centro, (TC), (público/privado) ejerce un efecto directo y positivo sobre la satisfacción de las familias, (RSF), de $Ed=.216$. Hemos de tener en cuenta, que en la muestra sólo hemos considerado centros privados-concertados, con un sistema de funcionamiento, en muchos aspectos, similar al público. Posiblemente, si se hubiesen considerado en la muestra centros privados, la influencia hubiese sido mayor.

La variable «status socioeconómico», (ASE), ejerce un efecto directo sobre el grado de satisfacción de los padres, (RSF), de $Ed=.422$. Del análisis causal se desprende que las familias de un «status socioeconómico» más alto tienen un nivel de satisfacción mayor hacia la escuela que las de menor «status». ¿Son estos resultados congruentes y no contradictorios con los encontrados a nivel de alumnos? El conocimiento de la realidad nos demuestra que no.

En primer lugar, las familias de status más alto, tienen más posibilidades de elegir el centro que desean, y, una vez elegido, si no responde a sus exigencias pueden elegir otro, no así las familias con menos posibilidades.

En segundo lugar, en cuanto a la diferente valoración sobre la escuela de las familias y los alumnos, es evidente que los padres y los alumnos valoran diferentes aspectos de la misma.

En tercer lugar, hemos de tener en cuenta que en la encuesta se pedía a los padres su propia opinión sobre la escuela, sin consultar con sus hijos.

Los recursos escolares, (D), ejercen un efecto total $Et=.277$, con un efecto indirecto $Ei=.101$ a través del Equipo Directivo, (OED), vía por la cual muchos padres se relacionan con la institución, confirmado por el valor total de OED con RSF, $Et=.281$. Pues, aunque los padres mantengan entrevistas con los tutores de sus hijos, quien proyecta a la comunidad «la imagen» de la escuela, es el Equipo Directivo de la misma.

Por otra parte, hemos de tener en cuenta que los problemas de mayor relieve como los de convivencia y los que se generan entre alumnos y tutores en los centros, son tratados por los padres con el Equipo Directivo en la mayoría de los casos.

CONCLUSIONES

A) Por lo que respecta a la influencia de las variables exógenas, podemos destacar:

- El factor con más influencia en la eficacia de los centros (rendimiento de los alumnos) es la dotación (recursos humanos y materiales) que posee la institución.
- El «status socioeconómico» influye en el nivel de satisfacción de alumnos y familias con la institución. Esta satisfacción es, por un lado, menor en los

alumnos de mayor «status» y, por otro, inversa a la de sus familias. Es decir, las familias de mayor «status» se encuentran más satisfechas con la institución escolar elegida.

- La estabilidad de los profesores en los centros influye positivamente en el funcionamiento del Equipo Directivo, Equipos de Profesores y en el trabajo de los profesores en las aulas y negativamente en el nivel de satisfacción de los alumnos.
- El centro privado-concertado influye positivamente en el funcionamiento del Equipo Directivo, en la satisfacción de las familias y en la disminución del índice de alumnos con problemas de aprendizaje.

B) En cuanto la influencia de las variables de proceso en la eficacia del centro podemos destacar que:

- El Equipo Directivo influye directa y positivamente en el funcionamiento del Consejo Escolar y, a través de éste, en el trabajo en equipo de los profesores y en la calidad de los procesos que se desarrollan en las aulas. Para ello, ha de promover una cultura organizativa caracterizada por la participación efectiva, la colaboración, la interdependencia, el autocontrol y la comunicación entre los profesores para implicar a todos en la marcha del centro y, por lo tanto en su transformación y mejora.
- El Consejo Escolar ejerce una influencia positiva en el funcionamiento en Equipo de los Profesores y, a través de éstos, en el clima del centro. Esta «Unidad Organizativa» se ha evidenciado como una estructura de gran importancia en la vida del centro cuando promueve metas y prácticas comunes, normas y valores consensuados, diferentes puntos de vista, discusiones y oposiciones. Pero todo esto, situado en un marco de «consenso» que todo el mundo suscribe, dentro del cual, los desacuerdos y/o conflictos se resuelven mediante el diálogo.
- El funcionamiento en Equipo de los Profesores, estudiando y analizando los problemas comunes de la enseñanza, buscando soluciones compartidas, explicitando, clarificando y consensuando lo que pretende conseguir y cómo conseguirlo, influye positivamente en el clima que se genera en el centro, en los procesos del aula y en la satisfacción del alumnado.
- El clima que se genera en el centro influye directa y positivamente en el trabajo de los profesores en las aulas y, a través de éste, en el rendimiento de los alumnos.

Un centro genera un clima positivo cuando es capaz de crear unas estructuras organizativas flexibles y operativas que permiten la participación de la Comunidad Educativa; el Equipo directivo ejerce un liderazgo efectivo; se promueven relaciones positivas entre sus miembros y se potencia el desarrollo y la satisfacción profesional del profesorado.

- El desarrollo de currícula de calidad exige de los profesores acciones diversificadas y complejas de planificación colaborativa, selección y organización

coherente de los contenidos, elaboración de programas de actividades diferenciadas, de utilización de estrategias específicas de enseñanza, de motivación y ayuda al alumnado, de elaboración de procedimientos de evaluación rigurosos, así como la creación de un ambiente de aula capaz de favorecer la implicación cognitiva y afectiva de los alumnos que favorezca la adquisición de aprendizajes valiosos y duraderos.

D) Acerca de las variables resultados hemos comprobado que:

- El rendimiento de los alumnos está influido por los recursos del centro, los procesos que se desarrollan en las aulas y el nivel de satisfacción de los propios alumnos.
 - En el nivel de satisfacción de los alumnos con la institución influyen el liderazgo del Equipo Directivo, el funcionamiento en Equipo de los Profesores y la calidad de los procesos que se desarrollan en las aulas.
 - El nivel de satisfacción de las familias está condicionado por el tipo de centro, la existencia y utilización adecuada de los recursos y el liderazgo del Equipo Directivo.
- El menor índice de alumnos con problemas de aprendizaje existe en los centros concertados y cuando hay profesorado con menor antigüedad.

Por último, aunque el patrón de relaciones planteadas en nuestro modelo, al ser validado empíricamente, ha resultado ajustarse bastante bien a los datos, lo hemos considerado como exploratorio, debido a las propias limitaciones de la investigación:

- Algunas de las variables consideradas han resultado poco explicadas por el modelo, siendo necesario identificar un mayor número de ellas para explicar la eficacia.
- La necesidad de profundizar en la conceptualización de las dimensiones constitutivas de los factores condicionadores de la eficacia, de manera que podamos llegar a un mejor conocimiento de los mismos.
- La posibilidad de que algunos de estos factores no mantengan entre sí una causalidad recursiva, si no más bien estructuras de relaciones recíprocas: Equipo Directivo y Consejo Escolar, Clima y Procesos de aulas, etc.

BIBLIOGRAFÍA

- ABALOS, J. y al. (1985): Statistical methods for selecting merit schools. En *Annual Meeting of American Educational Research Association*. New York.
- ANDERSON, C. (1982): The search for school climate: A review of the research. En *Review of Educational Research* nº 52, pp. 368-420.
- ANGUERA, M. T. (1985): *Manual de prácticas de observación*. Trillas.
- BALL, J. (1989): La micropolítica de la escuela. Paidós-MEC. Barcelona.
- BENSON, J. K. (1983): Organizations: A dialectical view. En FOSTER, W.: *Loose coupling revisite:*

- A critical view of Weick's contribution to educational administration.* Deakin Univ. Victoria, pp. 95-21.
- BERMAN, P. y McLAUGHLIN, M. V. (1976): Implementacion of educational innovation. *Educational Forum*.
- BISQUERRA, R. (1989): *Introducción conceptual al Análisis Multivariable*. Barcelona: PPU.
- BLUMBERG, A. y GREENFIELD, W. (1980): *The effective principal*. Allyn y Bacon, Boston.
- BOLAM, R. (1986): The national Developpe Centre for School Management Training. En Hoyle, E. y MCMahon, A.: *The mangement of Schools*. London, Kogan-Page.
- BORRELL, N. (1989): *Organizacion Escolar. Teoría sobre las corrientes científicas*. Ed. humanitas.
- BUSH, T. (1989): *A primer of LISREL Basic Aplications and Programming for Confirmatory Factor Analytic Models*. New York.
- CLARK, D. L.; LOTTO, T. A. y ASTUTO, T. A. (1984): Effective schools improvement: A comparative analysis of two lines of inquiry. *Educational Administration Quarterly*, vol. 20, 3, 41-68.
- CLIFF, N. (1983): Some cautions concerning the application of causal modeling methods. *Multivariate Behavioral Research*, 18, pp. 115-128.
- COHEN, M. D., MARCH, J. y OLSEN, J. (1972): A Garbage can Model of Organizational choice. *Administrative Science Quaterly*, 17, pp. 1-25.
- COLL, C. (1987): *Psicología y currículum: Aproximación psicopedagógica al currículum escolar*. Laia. Barcelona.
- CORBERT, H. D.; FIRESTONE, N. A. y ROSSMAN, G. B. (1987): Resistence to Planead Change and the Sacred in School Culture. *Educational Administration Quarterly*, 23, 4, pp. 353-390.
- CRANDALL, D. y otros (1982): Models of the School improvement process: Factors contributing to sucess. En *Annual Meeting of American Educational Research Association*. New York.
- CRONBACH, L. J. y al. (1972): *The dependability of Behavior Measurements: Theory of Generalizability for Scores and Profiles*. Nueva York. John Wiley.
- DEAL, T. y KENNEDY, A. (1983): *Corporate Cultures*. Reading Addison-Wesley.
- DEAL, T. (1985): The sinbolism os effective schools. *The Elementary School Journal*, 85, 5, pp. 601-620.
- DE MIGUEL, M. (1988): Modelos de investigación sobre las organizaciones educativas. *IV Seminario de modelos de investigación educativa*. Santiago.
- DE MIGUEL, M. (1988): Diseños de Investigación sobre las organizaciones educativas: Evaluación del cambio institucional. *Congreso de Pedagogía*. Alicante.
- ERIKCSEN, S. C. (1984): *The Essence of Good Teaching*. San Francisco, Jossey-Bass.
- ERICKSON, D. (1981): A New Strategy for school improvement. En *Momentum*, 12, 4.
- ERICKSON, F. (1987): Conceptions of School Culture: An overview. *Educational Administration Quartetly*. Vol, 23, 4.
- ERICKSON, F. y WILSON, J. (1982): Sights and Sounds of Life in Schools. Institute for Research on Teaching. *College of Education*. Michigan State University.
- EVERTSON, C. T. y otros (1983): Improving classroom management: An experiment in elementary school classrooms. *Elementary School Journal*, 84, 2, 173-188.
- EVERTSON, C. M. (1986): Observation as inquiry and method. En Wittock (Ed.): *Handbook of Research on Teaching*, 3ª ed. McMillan, New York.
- FUENTES, A. (1986): *Procesos funcionales y eficacia de la escuela. Un modelo causal. Tesis doctoral inédita*. Facultad de Ciencias de la Educación. Madrid.
- FULLAN, M. (1982): *The Meaning of Educational Change*. New York, Teacher College Press.
- FULLAN, M. (1986): The mangement of change. En Hoyle, E. y MacMahon, A.: *The management of School journal*. London. Kagan. Page.

- GLATTER, R. y otros (1988): *Understanding School management*. Philadelphia, Open University Press.
- GLATTER, R. (1990): La dirección como agente de innovación y cambio de los centros educativos. *Primer Congreso de Organización Escolar*. Barcelona.
- GOETZ, J. y LeCOMPTE, M. (1988): *Etnografía y diseño cualitativo en investigación educativa*. Ed. Morata. Madrid.
- GÓMEZ BENITO, J. (1986): *Los modelos causales como metodología de validez de constructo*. Barcelona: Alamex.
- GONZÁLEZ, M. T. (1989): La perspectiva interpretativa y la perspectiva crítica en organización. En *Martínez Moreno, Organizaciones Educativas*. UNED. Madrid.
- GOOD, T. L. y BROPHY, J. (1984): *Looking in classrooms*. 3ª ed. Harper and Row, Nueva York.
- GRAY, H. (1982): *The management of Educational Institutions: Theory, research and consultancy*. Lowes, Falmer Press.
- GRAY, J. L. (1985): *The School as an Organization*. Deanhouse. England.
- GREENFIELD, T. B. (1985): Theories of Educational Organizations: A Critical perspective. *International Encyclopedia of Education*. Oxford. Pergamon Press.
- HALL, G. E. y HORD, S. (1987): *Change in Schools: Facilitating The Process*. State University of New York Press, Albany, New York.
- HANDY, C. (1988): Cultural forces in schools. En GLATTER, R.: *Understanding School management*. Philadelphia, Open University Press.
- HANDY, Ch. B. (1984): *Taken for granted? Looking at schools as organizations*. Logman. York.
- HOPKINS, J. (1980): *Classroom Observation Project*. General Procedure for Reviews.
- HOPKINS, D. y WIDEEN, M. (1984): *Alternative perspectives on School improvement*. Falmer Press. Filadelfia.
- HUFF, S. y otros (1982): Principal differences, Excellence. En *School leadership and mangement*. Boston.
- JAMES, L.; MULAİK, S. A. y BRETT, J. M. (1982): *Causal analysis: assumptions, models and data*. Beverly Hills. Sage.
- JÖRESKOG, K. y SORBON, D. (1986): *LISREL VI: Analysis of linear structural relationships by maximun linkelihood. Instrumental variables and least squares methods*. University of Uppsala Fourt.
- LITTLE, J. W. (1982): The effective principal. *American Education*, 18, 7, pp. 38-43.
- LORTIE, D. (1969): The balance of control and autonomy in Elementary chool teaching. En ETZIO-NI, A.: *The semi-professionsand their organization*. Free Press. New York.
- MACKENZIE, D. E. (1983): Research for School Improvement: An Appraisal of same recent trends. En *Education Research*, 12, 4, pp. 5-17.
- MAHAFFY, J. (1989): *Collegial Support System*. Northwets Educational Laboratory, Portland. Oregon.
- MARCH, J. G. y OLSEN, J. (1976): *Ambiguity and choice in organizations*. Universitetsforlaget. Oslo.
- MARUYAMA, G. y WALBERG, M. (1982): Causal modeling. *Encyclopedia of Educational Research*, 5ª ed. Nueva York, AERA, MacMillan.
- McMAHON, H. y otros (1984): *Guidelines for review and internal development in schools (GRIDS)*. London. Longman.
- MEYER, J. W. y ROWAN, B. (1983): Institutionalized organization: Formal Structure as myth and ceremony. En *Meyer, J. y Scott, W.*

- MOOS, R. H. (1974): *Issues in Social Ecology*. National Press Books. Palo Alto.
- PEDHAZUR, E. J. (1982): *Multiple regression in behavioral research: explanation and prediction*. New York. 2º ed. Holt, Rinehart and Winston.
- PETERS, T. J. y WATERMAN, R. J. (1984): *En busca de la excelencia*. Barcelona. Folio.
- PONDY, L. R. (1983): *Union of rationaly and intuition in management action*. London, Jossey-Bass.
- PURKEY, S. C. y SMITH, M. S. (1983): Effective Schools. *Elementary school. Journal*, 4.
- RUNKEL, P. y SCHUMUCK, R. (1984): The Place of OD in school. En Hopkins y Wildeen: *Alternative Perspectives on School Improvement*. Falmer Press, London.
- SCHEIN, E. H. (1985): *Organizational culture and Leadership*. Jossey-Bass. San Francisco.
- SERGIOVANNI, T. J. (1987): *The principalship. A Reflective Practice Theory*. Allun and Bacon. Massachusetts.
- TYLER, W. (1986): *School organitation. A Sociological Perspective*. Croom Helm. London. New York.
- VISAUTA VINAGUA, A. (1986): *Técnicas de investigación social. Modelos causales*. Madrid, Hispano Europa.
- WALBERG, H. J. (1980): Educational climates. En Walberg: *Improving Educational Standars and Productivity: The Research Basis for Policy*. Berkeley. McCutchan.
- WEICK, K. E. (1976): Educational organizations as loosely coupled systems. *Administrative Science Quaterly*, 21, pp. 1-19.
- WEICK, K. E. (1979): *The social psichology of organizing*. Massachusetts. Addison-Wesley.
- WEINDLING, D. y EARLEY, P. (1988): How heads manage change. En Glatter, R. y otros: *Understanding School Management*, Milton Keynes, Open University Press.
- WOODS, P. (1987): *La escuela por dentro: La etnografía en la investigación educativa*. Madrid. M.E.C. Paidós.

ANÁLISIS Y VALORACIÓN DE PROGRAMAS PERTENECIENTES AL PLAN EXPERIMENTAL PARA LA REFORMA DE LA EDUCACIÓN INFANTIL

por

M^a Teresa Aguado Odina

Universidad Nacional de Educación a Distancia (UNED)

Facultad de Filosofía y Ciencias de la Educación

Departamento de Métodos de Investigación y Diagnóstico en Educación

RESUMEN

El artículo expone los aspectos fundamentales —justificación, plan experimental, conclusiones y recomendaciones— de la evaluación realizada de algunos programas educativos destinados a niños entre 4 y 6 años de edad, desarrollados en el marco del Plan Experimental para la Reforma de la Educación Infantil en España. El objetivo fundamental de la investigación ha sido analizar y valorar comparativamente proyectos de educación infantil pertenecientes al Plan Experimental y proyectos ajenos al mismo en orden a identificar las variables de entrada y proceso que diferencian entre programas reforma/no reforma y se relacionan con los resultados obtenidos por los alumnos; así como analizar los efectos diferenciales de los programas cuando se adopta como criterio el grado de consecución de los objetivos generales establecidos en la Reforma de la Educación Infantil.

ABSTRACT

The article presents the justification, experimental design, issues and conclusions about the evaluation of several Early Childhood Education programs (4 to 6 years old children) developed in the frame of the Experimental Plan for Educational Reform in Spain. The main objective of the research has been to analyze and evaluate different Reform and no reform educational projects in order to identify the input and process

variables that difference among the two groups of projects and to analyze the differential effects of the programs when we consider as criterion the general objectives established by the Educational Reform Project.

I. JUSTIFICACIÓN

La Reforma de la Educación Infantil en nuestro país fue iniciada por el MEC durante el curso 1985/86. En ese momento se seleccionaron los proyectos que han sido experimentados a lo largo de tres cursos en cada una de las Comunidades Autónomas que configuran el territorio MEC. Durante el curso 1989/90 comenzó la fase de difusión de los programas experimentales en áreas delimitadas del mapa escolar de cada provincia.

En diferentes momentos la Administración ha sometido a evaluación algunas dimensiones del proceso de implementación del Plan Experimental mediante el análisis de las valoraciones críticas realizadas por parte de los propios equipos docentes de los Centros, informes de los coordinadores provinciales y cuestionarios a los padres. No obstante, hasta el momento, no se ha llevado a cabo ninguna evaluación externa ni de las condiciones de aplicación y características de los proyectos desarrollados, ni de los resultados conseguidos durante la fase de experimentación.

El trabajo que aquí se presenta pretende aportar datos en esta línea investigadora y cubrir, así, un vacío en un tema que creemos de interés y actualidad. Junto a ésta, son varias las consideraciones que justifican nuestro interés:

— El escaso número de investigaciones rigurosas que analicen y evalúen —más allá de lo meramente descriptivo— el periodo correspondiente a la educación infantil, no sólo por lo que se refiere a nuestro país, sino, en general, en el contexto europeo.

— La política curricular de la Reforma educativa señala como una de las líneas directrices la de evaluar los diferentes elementos del sistema, elaborando criterios e instrumentos adecuados y comparando resultados en función de las diversas funciones desarrolladas (MEC, 1989).

— La reflexión personal y la actividad profesional más reciente nos han planteado una serie de interrogantes, tales como, ¿en qué se diferencian los modelos actualmente desarrollados en educación infantil?, ¿en qué medida los proyectos experimentales responden al enfoque planteado desde el MEC?, ¿qué resultados diferenciales se obtienen con unos u otros modelos?

Tres aspectos del trabajo creemos que deben ser destacados.

En primer lugar, el problema de investigación se plantea desde un *modelo diferencial* no centrado en los déficits sino en las diferencias entre los individuos objeto de estudio. Tales diferencias, una vez establecida su naturaleza e importancia educativa, deben ser tenidas en cuenta en el diseño de intervenciones educativas que optimicen el proceso y producto educativos.

En segundo lugar, se ha realizado un esfuerzo por tener en cuenta variables —tanto *estructurales como procesuales*— significativas en función de investigaciones previas (interacción padres/hijo, participación de los padres en el centro, estilo educativo familiar, satisfacción, experiencia y formación del profesorado, características del centro) y variables no examinadas en investigaciones anteriores (procesos curriculares en función del Diseño Curricular Base, actitud de padres y profesores ante la Reforma, funcionamiento de los equipos docentes).

Finalmente, se han aunado las dos perspectivas desde las que se ha abordado la evaluación de los programas de educación infantil —*transmisión cultural y desarrollo cognitivo* (fundamentalmente constructivista)— que, al mismo tiempo, coinciden con los dos enfoques que configuran preferentemente el modelo educativo propuesto para la reforma de la educación infantil (MEC, 1986/1989).

La perspectiva adoptada para afrontar el problema propuesto exige ir dando cumplimiento a las etapas habituales en todo proceso investigador: especificar objetivos, delimitar el estado actual de la cuestión, plantear hipótesis, elaborar y aplicar el diseño de investigación, establecer conclusiones.

2. OBJETIVOS DE LA INVESTIGACIÓN

Dos han sido los propósitos básicos de nuestra investigación:

1. Delimitar la Educación Infantil actual en cuanto a su conceptualización, modelos educativos desarrollados, estudios comparativos y evaluaciones de diferentes modelos y programas y proceso de reforma de esta etapa educativa en nuestro país.

2. Analizar y valorar comparativamente los diferentes proyectos de educación infantil desarrollados en el marco del Plan Experimental para la Reforma de la Educación Infantil y proyectos no pertenecientes a dicho Plan Experimental en orden a identificar las variables de entrada y proceso que diferencian entre programas reforma/no reforma y se relacionan con los resultados obtenidos por los alumnos; así como analizar los efectos diferenciales de los programas cuando se adopta como criterio el grado de consecución de los objetivos generales establecidos en la Reforma de la Educación Infantil (MEC, 1986/90).

Estos dos objetivos se complementan con otros dos que, consecuencia directa de los anteriores, son propios de las evaluaciones formativas orientadas hacia la toma de decisiones:

1. Proporcionar datos acerca de la situación actual del proceso de Reforma de la Educación Infantil y sobre las características estructurales y procesuales que lo diferencian de otras propuestas más tradicionales.

2. Mejorar el modelo educativo y los proyectos desarrollados en el ámbito de la Reforma de la Educación Infantil atendiendo a aquellas variables estructurales y procesuales identificadas como más influyentes en los resultados obtenidos por los alumnos.

3. PLAN EXPERIMENTAL

El plan experimental de investigación se orienta a dar respuesta a los tres últimos objetivos explicitados al inicio del presente trabajo.

3.1. Hipótesis

En correspondencia con los objetivos señalados al comienzo del trabajo y según los datos obtenidos acerca del estado actual del problema objeto de investigación, se plantean las siguientes hipótesis:

H.1. Los programas de Educación Infantil correspondientes a los dos últimos cursos de Educación Infantil pertenecientes al Plan Experimental y programas no pertenecientes al mismo se asocian a características diferenciales en las siguientes variables:

- Personales: escolaridad previa y programas de apoyo al alumno; formación, opiniones educativas, actitud ante la escuela y la reforma de los padres; satisfacción profesional, formación, experiencia y opiniones educativas de los profesores.
- Contextuales: ratio profesor/alumno, servicios de apoyo, infraestructura, estilo directivo, participación en programas y experiencias de innovación.
- Procesuales: estilo educativo familiar y relaciones de los padres con el Centro; modelo educativo, objetivos y grado de adecuación al DCB de los procesos desarrollados en clase.

H.2. Las variables que diferencian entre programas experimentales de la Reforma y programas no experimentales se relacionan con diferencias estadísticamente significativas en los resultados obtenidos por los alumnos al final de cada uno de los dos cursos de Educación Infantil estudiados en las variables:

- Cognoscitivas: capacidad intelectual general, desarrollo verbal, agudeza perceptiva, desarrollo conceptual, conocimiento físico, conocimiento lógico-matemático, capacidad para las matemáticas, las ciencias y la lectura, capacidad artística, conocimiento del entorno y creatividad.
- Afectivas: habilidades sociales, actitudes, hábitos de independencia, iniciativa, autoconfianza, persistencia, autocontrol y concentración.
- Psicomotricidad: motricidad fina y gruesa, coordinación motora.

H.3. Existen diferencias estadísticamente significativas en los resultados obtenidos a lo largo de los dos últimos cursos correspondientes al segundo ciclo de educación infantil por los alumnos de programas experimentales (GR) al ser comparados con los de los no experimentales (GNR) en el grado de consecución de los objetivos generales contenidos en el Proyecto para la Reforma de la Educación Infantil.

3.1. Estas diferencias se manifiestan a favor de los alumnos del GR en variables cognoscitivas —capacidad intelectual general, diversidad y complejidad de pensa-

miento, conocimiento lógico/matemático, desarrollo verbal, capacidad artística y creatividad— y afectivas —habilidades sociales, actitudes y hábitos de independencia, autoconfianza y cooperación, actitud hacia la escuela y autoconcepto— y desarrollo psicomotor.

3.2. Las diferencias se manifiestan a favor de los alumnos del GNR en variables cognoscitivas —capacidades intelectuales específicas (agudeza perceptiva, memoria, capacidad numérica, desarrollo conceptual y conocimiento del entorno), capacidades académicas— y afectivas —actitudes y hábitos de concentración, persistencia y autocontrol—.

H.4. Existen diferencias estadísticamente significativas en los resultados obtenidos a lo largo de los dos últimos cursos 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 ya mencionados.

4.1. Estas diferencias se manifiestan a favor de los alumnos de programas más adecuados al DCB en variables cognoscitivas —capacidad intelectual general, desarrollo verbal, complejidad y diversidad de pensamiento, conocimiento lógico-matemático, capacidades artísticas y creatividad—; afectivas —habilidades sociales, actitudes y hábitos de independencia, iniciativa, autoconfianza y cooperación— y desarrollo psicomotriz.

4.2. Las diferencias se manifiestan a favor de los alumnos de programas menos adecuados al DCB en variables cognoscitivas —capacidades intelectuales específicas (agudeza perceptiva, memoria, capacidad numérica, desarrollo conceptual y conocimiento del entorno), capacidades académicas— y afectivas —actitudes y hábitos de concentración, persistencia y autocontrol—.

H.5. Existen diferencias estadísticamente significativas en los resultados obtenidos a lo largo de los dos últimos cursos correspondientes al segundo ciclo de educación infantil por los alumnos de programas que responden a diferentes modelos educativos en los objetivos generales ya mencionados.

5.1. Estas diferencias se manifiestan a favor de los modelos cognoscitivos en variables cognitivas —capacidad intelectual general, desarrollo verbal, complejidad y diversidad de pensamiento, conocimiento lógico-matemático, capacidades artísticas y creatividad—; afectivas —habilidades sociales, actitudes y hábitos de independencia, iniciativa y autoconfianza—.

5.2. Las diferencias se presentan a favor de los alumnos de modelos preacadémicos en variables cognoscitivas —capacidades intelectuales específicas (agudeza perceptiva, memoria, capacidad numérica, desarrollo conceptual y conocimiento del entorno); capacidades académicas—, y afectivas —hábitos de concentración y autocontrol—.

5.3. Las diferencias se manifiestan a favor de los modelos maduracionistas en variables afectivas —actitud hacia la escuela y autoconcepto— y desarrollo psicomotor.

H.6. Existen diferencias estadísticamente significativas en los resultados obtenidos a lo largo de los dos últimos cursos de educación infantil a favor de los alumnos participantes en el programa desarrollado por los padres en el hogar al ser comparados con los de los alumnos no participantes en tal programa. Estas diferencias se manifiestan en las siguientes variables: desarrollo verbal, agudeza perceptiva, memoria, desarrollo conceptual, conocimiento del entorno, capacidad para la lectura y creatividad.

3.2. Diseño

Se ha acudido a 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 cuándo 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 toman medidas en momentos dados del proceso— y longitudinal de tratamiento —se van a realizar 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 procesuales—.

Como todo estudio longitudinal permitirá realizar estimaciones iniciales como línea base que hagan posible valorar el cambio individual y la incidencia del tratamiento (De Miguel, 1988); observar la magnitud y persistencia de los cambios a lo largo del periodo de intervención (Rogosa, 1980); superar algunas limitaciones metodológicas al controlar los efectos de la maduración y tener en cuenta variables procesuales y contextuales relevantes (Alkin, 1990; Cronbach, 1980; Weir, 1975).

Para dar respuesta a la última hipótesis referida a los efectos de un programa de intervención diseñado para ser implementado por los padres en el hogar durante los dos cursos académicos del segundo ciclo de educación infantil, el diseño central se complementa con otro de carácter experimental. El grupo experimental está formado por los alumnos de los dos centros situados en el extrarradio en los que se implementará el programa de intervención. El grupo control lo constituyen el resto de los alumnos de la muestra no participantes en dicho programa.

3.3. Variables

La selección de las variables a analizar se ha realizado teniendo en cuenta los siguientes criterios:

1. Ser significativas para los objetivos que planteamos según la evidencia aportada por las investigaciones previas revisadas.
2. Considerar variables independientes tanto estructurales —de entrada— como

procesuales —procesos reales observados en la aplicación práctica de los proyectos— (Weir, 1975; Cronbach, 1980; Alkin, 1990).

3. Tener en cuenta el enfoque predominantemente cognitivo del modelo de educación infantil propuesto en los documentos del proyecto de Reforma: «Anteproyecto de Marco Curricular» (MEC, 1986); «Proyecto de Reforma de la Enseñanza» (MEC, 1987); «Diseño Curricular Base» (MEC, 1989), sin obviar variables propias del enfoque de la transmisión cultural.

4. Adoptar como variables dependientes o criterio medidas de los objetivos generales establecidos por el Proyecto de reforma para la Educación Infantil tal y como se exponen en los documentos mencionados.

3.4. Instrumentación

Los instrumentos utilizados para la medición de las variables consideradas se agrupan para su exposición y análisis en dos categorías: pruebas standarizadas e instrumentos elaborados «ad hoc». En la selección y elaboración de unos y otros se han tenido en cuenta las siguientes consideraciones:

1. Ser adecuados a las hipótesis planteadas en la investigación.
2. Garantías suficientes de validez y fiabilidad y/o utilización eficaz en investigaciones previas.
3. Permitir el contraste de datos mediante la utilización de diversas fuentes - triangulación - como en el caso de variables medidas mediante la aplicación de pruebas estandarizadas, observación en clase y observación en el hogar.
4. Ajustados a los recursos humanos y materiales disponibles.

Junto a instrumentos orientados a proporcionar medidas de ejecución en tareas estructuradas, se han seleccionado y/o elaborado pruebas basadas en la medición observacional, es decir, «en el registro sistemático de conductas o contextos de forma que se obtengan descripciones y medidas cuantitativas de individuos, grupos o ambientes». (Goodwin y Driscoll, 1984:110).

La utilización de este tipo de instrumentos se justifica por tres razones fundamentales: 1) el repertorio de pruebas estandarizadas específicas para la etapa 0/6 años es muy limitado por lo que se hace preciso elaborar o adaptar instrumentos adecuados a los objetivos de la investigación; 2) la observación directa es especialmente adecuada para observar conductas de niños menores de seis años dado su limitado repertorio de respuestas - especialmente verbales - (Walker, 1973) y su «imperturbabilidad» ante la observación de otros (Wright, 1967); 3) la observación directa es especialmente adecuada y, a menudo insustituible, para registrar procesos tales como actividades, estrategias e interacciones desarrollados en la clase; y 4) Se evitan los sesgos debidos al aprendizaje propios de las pruebas estructuradas cuando son utilizadas en aplicaciones repetidas en diseños longitudinales.

3.5. Muestra

La muestra final está constituida por ciento ochenta y cuatro alumnos de Colegios Públicos situados en Zaragoza capital que han cursado los dos últimos años de educación infantil durante el bienio 1988/90, habiendo permanecido durante los dos cursos en el mismo grupo-clase. De ellos, ciento treinta pertenecen a cuatro Centros integrados en el Plan Experimental para la Reforma de la Educación Infantil y cincuenta y cuatro a dos Centros no Experimentales. Como ya se ha señalado al exponer los criterios adoptados en la selección de la muestra, en algunos centros ha sido preciso utilizar más de un grupo-clase atendiendo a la previsible mortalidad que podía producirse durante los dos años del estudio.

3.6. Análisis Estadísticos

Una vez preparados los datos en formato informático, se ha procedido a la realización de los correspondientes análisis estadísticos mediante la aplicación del paquete estadístico SPSS/PC+ siguiendo los siguientes pasos:

A) En un primer momento se han obtenido los descriptivos correspondientes a cada una de las variables analizadas: media, error standar, mediana, moda, desviación estándar, varianza, curtosis y simetría, rango, valores máximo y mínimo, casos válidos.

B) Para la comprobación de la hipótesis primera se ha procedido a la aplicación de medidas de asociación. En total se han efectuado sesenta tablas cruzadas entre la variable CENREFOR («grupo reforma» o experimental y grupo «no reforma» o control) y cada una de las variables no continuas especificadas en las tres subhipótesis correspondientes. La prueba estadística aplicada ha sido chi-cuadrado (X) mediante la cual se ha sometido a comprobación la independencia entre cada par de variables analizadas en cada tabla. En el caso de variables continuas, se ha aplicado el correspondiente análisis de varianza entre los GR/GNR para cada una de las variables analizadas. En total se han efectuado treinta ANOVAs. La homogeneidad de varianzas exigida por esta prueba se ha comprobado mediante las pruebas C de Cochran y la F de Bartlett-Box.

C) La comprobación de la hipótesis segunda ha supuesto la aplicación de medidas correlacionales. Así, se ha procedido a la realización de los correspondientes análisis de varianza, tablas cruzadas (prueba chi-cuadrado) y coeficiente de correlación de Pearson para establecer las posibles correlaciones entre variables categoriales/continuas, categoriales/categoriales y continuas/continuas, respectivamente. En total se han realizado 427 análisis de varianza, 149 tablas cruzadas y 280 coeficientes de correlación.

D) La comprobación de la hipótesis tercera ha supuesto la aplicación de medidas inferenciales. Así, se ha procedido a realizar los correspondientes contrastes entre el GR y el GNR en orden a establecer la existencia o no de diferencias estadísticamen-

te significativas entre ambos grupos en cada una de las variables criterio establecidas y para cada una de las seis mediciones efectuadas. Para ello se ha procedido a la realización de los siguientes análisis:

a) Análisis de varianza para cada variable continua en la primera medida efectuada con objeto de establecer las diferencias iniciales entre los dos grupos (GR/GNR). En total se han efectuado cuarenta y un ANOVAs.

b) Tablas cruzadas y aplicación de las mismas pruebas especificadas en la hipótesis primera para cada una de las medidas correspondientes a las variables no continuas. En total se han realizado treinta y una tablas cruzadas.

c) Análisis de covarianza para medidas repetidas. Con objeto de realizar un análisis temporal y establecer el momento en que aparecen las posibles diferencias entre los grupos en comparación, se han realizado diversos análisis tomando como covariable diferentes medidas de la misma variable.

1. En el caso de variables medidas en seis ocasiones a lo largo de los dos cursos, los análisis de covarianza se han aplicado para las medidas:

— 2^a, 3^a, 4^a, 5^a y 6^a mediciones de cada variable tomando como covariable la primera medida efectuada (considerada como pretest por haberse efectuado en la segunda semana del primer curso).

— 3^a medición efectuada (final del primer curso) tomando como covariable la 2^a medida.

— 5^a y 6^a tomando como covariable la 4^a (efectuada a comienzos del segundo curso).

— 6^a tomando como covariable la 5^a medida.

2. En el caso de variables medidas en dos ocasiones (al comienzo del primer curso y a finales del segundo) se ha aplicado el análisis de covarianza para la segunda medida de cada variable tomando como covariable la primera medida efectuada (pretest). En el caso de valores de F significativos a un nivel del 1% o el 5% y, por tanto, de diferencias estadísticamente significativas entre los grupos comparados en la variable analizada, se ha procedido a la aplicación de la prueba Scheffé para establecer entre qué grupos se manifestaban diferencias significativas a un nivel mínimo del 5%.

En total se han efectuado doscientos ocho análisis de covarianza.

E) Para la comprobación de las hipótesis cuarta y quinta se han realizado los mismos análisis especificados en el apartado anterior. En este caso con objeto de comparar los resultados obtenidos en las variables criterio por los grupos establecidos atendiendo al grado de adecuación al DCB (H.4) y el modelo educativo desarrollado en los programas (H.5).

F) La comprobación de la H.6. ha supuesto la aplicación de los correspondientes análisis de varianza y covarianza para cada una de las variables criterio consideradas. En total se han realizado setenta análisis.

G) Finalmente, se han realizado dos análisis discriminantes con objeto de establecer variables de entrada y proceso, en un caso, y variables criterio, en otro, que discriminan entre los grupos reforma y no reforma estudiados. El criterio de selec-

ción de variables utilizado es el lambda de Wilks. La hipótesis de igualdad entre matrices varianzas-covarianzas se ha comprobado mediante la prueba M de Box. En el primero de los análisis se han introducido dieciséis variables —de entrada y proceso— predictoras. En el segundo, trece variables criterio correspondientes a resultados obtenidos por los alumnos en la última medición efectuada al final del segundo curso del estudio.

4. CONCLUSIONES

A partir de los resultados obtenidos en la comprobación de cada una de las hipótesis planteadas, se han establecido las conclusiones que a continuación se presentan brevemente.

A) Por lo que se refiere a la primera hipótesis planteada en relación a las *diferencias entre los programas del grupo reforma y no reforma (GR/GNR)*, los resultados obtenidos al comparar los programas pertenecientes al Plan Experimental para la Reforma de la Educación Infantil y programas no pertenecientes al mismo nos permiten concluir que:

— Por lo que se refiere a las *variables personales* analizadas, las diferencias entre ambos grupos se manifiestan en escolaridad previa de los alumnos, actitud hacia la escuela y formación educativa de los padres al final del segundo curso. Por lo que se refiere al profesorado, las diferencias se constatan en objetivos educativos priorizados, grado de satisfacción profesional, titulación, especialización, actividades de formación permanente, experiencia y actitud hacia la reforma y, finalmente, objetivos priorizados para la etapa de educación infantil.

Los alumnos de programas desarrollados en el Plan Experimental para la Reforma de la Educación Infantil presentan una mayor tasa de escolarización a los tres años. Los padres de los alumnos de estos programas muestran con respecto a los del grupo no perteneciente al Plan Experimental al final de los dos cursos estudiados, una actitud hacia la escuela más positiva, superior nivel de lecturas y preparación en temas educativos, mayor participación en actividades de formación, mayor ajuste entre los objetivos que consideran prioritarios en educación infantil y los propuestos por el programa al que asiste su hijo. El profesorado de estos programas manifiesta un mayor grado de satisfacción profesional y participación en actividades de formación permanente de forma continua y focalizada en temas relacionados con los objetivos del programa.

No se han encontrado diferencias entre los grupos reforma y no reforma en lo que se refiere a implementación de programas para alumnos con necesidades específicas y generación en los padres de actitudes más positivas hacia la reforma de la educación infantil.

— Atendiendo a las *variables contextuales*, los programas correspondientes al Plan Experimental analizados se desarrollan en centros completos de EGB con más

de cuatro unidades de preescolar y una ratio superior a veinte alumnos por profesor. Los servicios de apoyo disponibles son los Equipos Multiprofesionales de zona y equipos específicos de Atención Temprana. No se dispone de profesores de apoyo que actúen de forma directa en el aula. La infraestructura de los centros es adecuada, el estilo directivo es democrático y participan en otros programas de experimentación e innovación de forma continuada.

— En lo que concierne a las *variables procesuales*, se distingue entre las referidas al ámbito familiar y el escolar. Así, en el ámbito familiar, los efectos diferenciales de los programas de reforma se manifiestan preferentemente en las relaciones que los padres mantienen con el centro y los profesores, puesto que, al final del segundo curso, los contactos son más frecuentes que en el grupo no reforma y los temas abordados se refieren no sólo al comportamiento y progresos académicos del alumno sino también a informaciones y recomendaciones sobre el programa educativo global desarrollado.

En cuanto a los procesos desarrollados en la clase, los programas del grupo reforma responden a modelos educativos diversos entre los que se encuentran representados los tres modelos básicos —maduracionista, preacadémico y desarrollo cognitivo—. Los objetivos sobre los que se focalizan los programas abarcan las tres áreas establecidas —cognoscitiva/verbal, afectiva y psicomotriz— de forma equilibrada. En todos los casos, los procesos observados en clase —medidos mediante la escala elaborada «ad hoc» (ver anexo nº 6)— presentan un mayor grado de adecuación a los principios organizativo/didácticos establecidos en el Diseño Curricular Base que los programas no reforma.

Por su parte, los programas no pertenecientes al grupo reforma responden a modelos educativos preferentemente preacadémicos, sus objetivos se focalizan en el desarrollo de habilidades y hábitos de preparación para la escuela y el grado de adecuación al DCB de los procesos observados en clase es menor.

— Las variables que mejor discriminan entre los programas correspondientes a los grupos reforma y no reforma son la conducta del profesor —medida mediante la escala para la identificación del grado de adecuación al DCB—, infraestructura del Centro, variedad y organización del espacio y los materiales, ratio profesor/alumno y temas abordados en los contactos de los padres con los profesores.

B) Atendiendo a la posible *relación entre variables que diferencian entre programas reforma / no reforma y resultados de los alumnos (hipótesis segunda)*, los resultados obtenidos permiten concluir que:

— Por lo que respecta a las *variables personales* analizadas, la escolaridad previa del alumno a los tres años se relaciona con mejores resultados al final del segundo curso en capacidades generales básicas —agudeza perceptiva—, desarrollo conceptual, diversidad de pensamiento y conducta independiente. De especial interés es constatar que son los alumnos que han asistido a escuelas infantiles —mayoritariamente en el mismo centro— los que obtienen mejores resultados en variables directamente relacionadas con las metas del programa, esto es, capaci-

dad para las ciencias y las matemáticas, capacidad artística, creatividad, habilidades sociales en la escuela y motricidad. Ninguna de las variables personales analizadas presenta relaciones significativas con los resultados obtenidos por los alumnos en variables del ámbito afectivo.

— De las *variables contextuales* estudiadas, tres establecen diferencias significativas entre los programas reforma y no reforma: ratio profesor/alumno, equipos de apoyo e infraestructura del centro. Cuando se considera la ratio, es el grupo con una proporción entre 20/30 alumnos por profesor el que obtienen mejores resultados en habilidades sociales, capacidad para las matemáticas, capacidades artísticas y desarrollo motriz. Los grupos con ratio más alta obtienen mejores puntuaciones en hábitos de persistencia y autocontrol.

Se constatan, asimismo, los efectos positivos de la actuación de los Equipos de Atención Temprana. Sus funciones van desde proporcionar orientaciones y recursos didácticos a los profesores hasta diseñar programas individualizados destinados a superar déficits específicos. Sus efectos se evidencian, especialmente, durante el primer curso estudiado, en cuanto al desarrollo de la capacidad intelectual general, agudeza perceptiva y motricidad. Durante el segundo curso, los efectos positivos se manifiestan en las variables diversidad de pensamiento y motricidad.

El grupo con una infraestructura del centro más adecuada presenta mejores resultados en variables cognoscitivas y motrices: capacidad intelectual general, diversidad de pensamiento, tareas de clasificación y seriación, capacidad para las ciencias, creatividad y motricidad.

— Atendiendo a las *variables procesuales familiares* que diferencian los grupos reforma y no reforma, se constata cómo es el grupo cuyos padres mantienen más frecuentes contactos con los profesores en los que abordan informaciones y recomendaciones referidas al programa educativo global el que presenta superiores resultados en variables cognoscitivas —capacidad para las matemáticas, lectura y plástica, cuando se atiende a la frecuencia de los contactos al comienzo de la etapa— y afectivas —habilidades sociales, hábitos de independencia, iniciativa, autoconfianza y persistencia, tanto al considerar la frecuencia de los contactos padres/profesores al comienzo como al final de preescolar—. Es relevante señalar la existencia de diferencias significativas a favor del grupo de alumnos cuyos padres mantienen algún tipo de contacto con los profesores respecto al grupo que no los mantiene en hábitos y actitudes de independencia, iniciativa y autoconfianza. Esto podría explicarse, en parte, al cambio en la percepción y valoración que los padres hacen de la conducta de su hijo.

— Por último, en cuanto a las *variables procesuales escolares*, la mayor adecuación de los procesos observados en clase a los principios organizativo/didácticos establecidos en el Diseño Curricular Base (DCB) se relacionan positivamente con superiores resultados en las siguientes variables: diversidad y complejidad de pensamiento, tareas de clasificación, capacidades artísticas, capacidad para las ciencias, comprensión de la información procedente del entorno y hábitos de independencia y autoconfianza.

La relación encontrada entre algunas de las variables que diferencian entre GR/GNR y los resultados de los alumnos al final de la etapa de educación infantil son especialmente significativos por cuanto que se trata de variables sobre las que es posible actuar y modular en orden a mejorar el grado de consecución de los objetivos propuestos.

C) En cuanto a los *efectos diferenciales de los programas de los grupos reforma y no reforma (hipótesis tercera)*, los datos obtenidos nos permiten concluir que:

— En las *variables cognoscitivo-verbales*, los programas del grupo reforma muestran una mayor eficacia que los programas del grupo no reforma en el desarrollo de la capacidad intelectual general medida mediante observación directa en la escuela, agudeza perceptiva, complejidad y diversidad de pensamiento, tareas de seriación, concepto de número, capacidad para la plástica, matemáticas y ciencias. Los programas del grupo no reforma son más eficaces en el desarrollo de capacidades medidas mediante pruebas altamente estructuradas —capacidad intelectual general, agudeza perceptiva, capacidad numérica y memoria— desarrollo conceptual, tareas de clasificación y comprensión de la información del entorno. No se constatan diferencias entre los programas de uno y otro grupo en desarrollo verbal, capacidad para la lectura y la música y creatividad.

Los efectos diferenciales a favor de unos u otros programas se manifiestan de forma casi exclusiva durante el segundo curso del estudio y, preferentemente, cuando la covariable adoptada para realizar los contrastes es el pretest aplicado al inicio del primer curso. Junto a esto, se constata como diferencias iniciales a favor de uno de los dos grupos desaparecen o cambian de signo al final de los dos cursos durante los que se ha realizado el seguimiento. Esto confirmaría la evidencia acerca de la alta maleabilidad de las aptitudes y habilidades del grupo de edad estudiado y, por tanto, la decisiva importancia de la intervención educativa en esta etapa.

— En lo relativo a las variables del *área afectiva*, no se constatan diferencias en ninguna de las variables analizadas: habilidades sociales en la familia y la escuela, autoconcepto, actitud hacia la escuela, hábitos de independencia, autoconfianza, iniciativa, autocontrol, concentración y persistencia.

— Por último, al considerar las variables del *ámbito psicomotor*, no se han constatado diferencias significativas entre los dos grupos en los resultados correspondientes a las tres pruebas utilizadas en ninguno de los momentos en que se ha procedido a su aplicación.

— Las variables criterio que, medidas al final del segundo curso, mejor discriminan entre los programas reforma y no reforma son y por este orden: agudeza perceptiva y capacidad intelectual general medidas mediante test estandarizado, actividades de clasificación, capacidad para las ciencias, capacidad intelectual medida mediante observación directa en la escuela, capacidad para la plástica, comprensión de la información procedente del entorno y desarrollo conceptual.

En líneas generales, se constata que las diferencias entre ambos grupos de

programas se presentan únicamente en el área cognoscitiva. Los programas del GR son superiores en el desarrollo de capacidades e intereses intelectuales y académicos —matemáticas y ciencias— medidos mediante observación directa en clase; transición al periodo de operaciones concretas —complejidad y diversidad de pensamiento y conocimiento lógico/matemático— y capacidad para la plástica. Por su parte, los programas del GNR son más eficaces en el desarrollo de la capacidad intelectual general y capacidades específicas medidas mediante pruebas de ejecución altamente estructuradas; así como en tareas de clasificación y en comprensión de la información del entorno.

D) Por lo que respecta a los *efectos diferenciales de los programas agrupados en función de su grado de adecuación al DCB (hipótesis cuarta)*, los datos obtenidos nos permiten concluir que:

— 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.

No se han constatado diferencias entre los dos grupos establecidos en el desarrollo de la capacidad intelectual general medida mediante observación directa, desarrollo verbal, capacidad para las ciencias y la lectura, información sobre el entorno y creatividad medida en la escuela.

— 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.

— 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.

E) Al considerar los *efectos diferenciales de los programas en función del*

modelo educativo desarrollado (hipótesis quinta), los resultados obtenidos permiten establecer las siguientes conclusiones:

— En el *área cognoscitivo/verbal*, el modelo equilibrado —caracterizado por presentar componentes propios de los tres modelos básicos considerados— obtiene superiores resultados respecto a los otros tres modelos considerados, al final del primer curso, en el desarrollo de la capacidad intelectual general medida mediante prueba altamente estructurada y capacidad perceptiva. Al final del segundo curso, su eficacia se manifiesta en superiores resultados en capacidad perceptiva y diversidad de pensamiento.

El modelo preacadémico/maduracionista —en el que predominan elementos del modelo preacadémico junto con características del maduracionista— obtiene superiores resultados durante el primer curso en desarrollo conceptual y, al final del segundo curso, en capacidad para las matemáticas y la música.

El modelo cognoscitivo/maduracionista —con una alta puntuación en las dimensiones propias del modelo cognoscitivo aunque con algunas características del enfoque maduracionista— presenta mejores resultados a lo largo del segundo curso en tareas de conocimiento físico —diversidad y complejidad de pensamiento—, actividades de clasificación, seriación y concepto de número; capacidad para las ciencias y la lectura y creatividad medida mediante observación directa en la escuela y en la familia.

La superior eficacia del modelo preacadémico —único representante «puro» de uno de los modelos básicos considerados— se manifiesta al final del segundo curso en una serie de capacidades medidas mediante prueba altamente estructurada —capacidad intelectual general, desarrollo verbal, agudeza perceptiva, capacidad numérica, memoria, desarrollo conceptual, comprensión de la información del entorno— y capacidad para la plástica.

— Si atendemos a las *variables afectivas* seleccionadas, se pone de manifiesto una mayor efectividad del modelo cognoscitivo/maduracionista en el desarrollo de hábitos y actitudes de autoconfianza, iniciativa, persistencia y autocontrol, y del modelo preacadémico en habilidades sociales en la escuela.

— En lo que concierne al *área psicomotriz*, es el modelo preacadémico el que obtiene mejores resultados durante las tres mediciones efectuadas a lo largo del segundo curso en el desarrollo de capacidades motrices cuando se miden mediante prueba altamente estructurada y estandarizada. Los modelos que incluyen componentes maduracionistas relevantes —preacadémico/maduracionista y cognoscitivo/maduracionista— se muestran más eficaces a lo largo de los dos cursos en el desarrollo de capacidades motrices medidas mediante observación directa en la escuela y en la familia, respectivamente.

E) Por último, al evaluar la *eficacia del programa complementario aplicado por los padres (hipótesis sexta)*, se concluye que:

— Se han encontrado diferencias significativas entre ambos grupos en desarrollo verbal, información procedente del entorno y creatividad, pero sólo en esta última variable son favorables al grupo participante en el programa.

— El grupo participante en el programa obtiene superiores resultados en creatividad medida mediante observación directa en la familia. Esto, refleja efectos positivos indirectos del programa, ya que la interacción establecida durante la realización de las actividades del programa permite a los padres atender a capacidades y conductas de su hijo hasta entonces ignoradas.

— En la variable «capacidad para la lectura», si bien las diferencias entre los dos grupos considerados no son estadísticamente significativas, se observa a lo largo de las sucesivas mediciones efectuadas durante los dos cursos de aplicación, una superior y más continua mejora de las puntuaciones medias del grupo participante en el programa con respecto a las del grupo no participante.

— Por último, es importante señalar la valoración altamente positiva que, tanto padres como profesores, han realizado del programa. Los primeros lo consideran una actividad sencilla que les permite interactuar con su hijo y favorecer su desarrollo. Los segundos estiman que, al margen de sus potenciales efectos positivos en el desarrollo del niño, es una estrategia adecuada para iniciar y estimular la participación de los padres en el proceso educativo de su hijo. Todos los profesores y padres participantes en el programa han continuado desarrollándolo en los cursos sucesivos posteriores a nuestra investigación.

5. RECOMENDACIONES

Atendiendo al objetivo ya señalado de mejorar el diseño e implementación de los programas desarrollados en el ámbito de la Reforma de la Educación Infantil y considerando aquellas variables estructurales y procesuales identificadas como más influyentes en los resultados obtenidos por los alumnos, se plantean una serie de recomendaciones orientadas a potenciar objetivos y acciones ya contenidos en el Proyecto para la Reforma de la Educación Infantil cuya relevancia y eficacia han sido puestos de manifiesto en este estudio, así como algunas orientaciones destinadas a los equipos docentes implicados en el diseño e implementación de programas específicos.

Entre las primeras, se señalan:

— Extender la escolaridad a los tres años en todos los centros públicos y privados que desarrollan programas de educación infantil.

— Atención diferencial a alumnos, no pertenecientes al plan de integración escolar, con necesidades específicas tanto de aprendizaje como características.

— Planificar y desarrollar acciones de información a los padres sobre los nuevos planteamientos educativos introducidos por el Proyecto de Reforma.

— Fomento de la participación de los padres en el programa educativo. Esto supone tanto la planificación de actividades de formación para los padres —grupos de trabajo, charlas, lecturas— como participación directa en programación y desarrollo de actividades con los niños en la escuela y en el hogar.

— Facilitar la formación permanente de los profesores, tanto individual como en

grupos de trabajo que fomenten la reflexión y el debate. Esto supone la adopción de medidas administrativas que permitan al profesorado implicarse en actividades de formación sin excesiva sobrecarga de trabajo y constituir equipos estables y cohesionados.

— Mejorar la infraestructura de los centros tanto en lo que se refiere a espacios y materiales —diversificados y adaptados a las necesidades del programa— como a recursos humanos —profesorado y equipos de apoyo—.

— Elaboración de materiales de apoyo que provoquen la reflexión y orienten la práctica. En ellos se contendrían tanto principios generales acerca de dimensiones relevantes del proceso educativo como ejemplificaciones de acciones concretas desarrolladas por los equipos docentes de los centros y evaluaciones de los resultados obtenidos.

— Implicación más directa en el asesoramiento y seguimiento de la planificación, implementación y evaluación de los programas experimentales de los Centros de Profesores a través de sus asesores y especialistas, de las Direcciones Provinciales a través de sus inspectores técnicos y coordinadores, y de los Departamentos universitarios a través de sus equipos docentes e investigadores.

Los equipos docentes responsables del diseño e implementación de programas de Educación Infantil deberían tener en cuenta que:

— Ningún modelo o programa se ha mostrado como superior a cualquier otro. Diferentes modelos educativos, ejemplificados por diferentes procesos, ofrecen diferentes ventajas. Así, el diseño y planificación del programa habrá de realizarse atendiendo a las metas y objetivos priorizados.

— El diseño curricular debería derivarse de la reflexión personal, consulta de materiales, debate e intercambio de experiencias y tener en cuenta las características personales, —procedencia sociocultural, edad y escolaridad previa— de los alumnos —y el contexto, tipo de centro, infraestructura, servicios de apoyo, ratio profesor/alumno— en que se va a aplicar.

— El programa debe implementarse de forma que se logre el mayor grado de ajuste posible entre el planteamiento teórico, objetivos priorizados y procesos —conducta del profesor, contenidos, características y organización de espacios y materiales, temporalización, participación de los padres, funcionamiento del equipo docente, evaluación de los resultados—.

— No existe una relación lineal y directa entre las dimensiones relevantes del programa educativo —personales, contextuales y procesuales— y los resultados de los alumnos. La evidencia acumulada permite establecer que a partir de un mínimo indispensable en las variables que identifican a los programas más eficaces, las características y calidad del proceso son más importantes que la cantidad —de profesorado y apoyos, contenidos, materiales, espacios, tiempo, actividades, formación de los participantes, etc.—.

REFERENCIAS BIBLIOGRÁFICAS

- ALKIN, M. C. (1990): «Curriculum evaluation models», en H. K. WALBERG, G. D. HAERTEL (eds.): *The International Encyclopedia of Educational Evaluation*. Pergamon Press. Oxford, 166-168.
- BARNETT, P. (1985): *The Perry School Program and its long-term effects: A Benefit - Cost Analysis*. High/Scope Press. Ypsilanti, Michigan.
- BIBER, B. (1984): *Early Education and Psychological Development*. Yale University Press. New Haven. London.
- BLOOM, B. S. (1964): *Stability and change in human characteristics*. John Wiley and Sons. New York.
- CLARK - STEWART, A. (1982): *Guarderías y cuidado infantil*. Morata. Madrid.
- CRONBACH, L. J. (1980): *Toward reform of program evaluation*. Jossey-Bass. San Francisco.
- DECKER, C. A. y DECKER, J. R. (1988): *Planning and Administering Early Childhood Education*. Merrill Publishing Company. Columbus. Ohio.
- DE KETELE, J. M. (1984): *Observar para educar. Observación y evaluación en la práctica educativa*. Aprendizaje-Visor. Madrid.
- DE MIGUEL, M. (1988): *Preescolarización y rendimiento académico: un estudio longitudinal de las variables psicosociales a lo largo de la E.G.B.* C.I.D.E. Madrid.
- GOODWIN, W. L. y DRISCOLL, L. A. (1984): *Measurement and Evaluation in Early Childhood Education*. Jossey-Bass. San Francisco.
- KAMII, C. y DE VRIES, R. (1985): *La teoría de Piaget y la educación preescolar*. Aprendizaje - Visor. Madrid.
- MINISTERIO DE EDUCACIÓN (1986/90): *Proyecto para la Reforma de la Educación Infantil*. Madrid.
- PETERS, D. L.; NEISWORTH, J. T. y YAWKEY, T. D. (1985): *Early Childhood Education: From Theory to Practice*. Brooks/Cole Publishers. Monterey.
- STUFFLEBEAM, D. L. y SHINKFIELD, A. J. (1987): *Evaluación sistemática. Guía teórica y práctica*. Paidós/MEC. Madrid.
- WALKER, D. K. (1973): *Socioemotional Measures for Preschool and Kindergarten Children*. Jossey-Bass. San Francisco.
- WEDELL - MONNING, J. y Mc NEIL, J. (1980): *National Evaluation of Head Start Education Services and Basic Educative Skills* Research Corporation, Durham, N.C.
- WRIGHT, H. F. (1967): *Recording and Analyzing Child Behavior*. Harper and Row. New York.

MODELO PARA UNA AUTOEVALUACIÓN MODULAR EN UNA INSTITUCIÓN EDUCATIVA (B.A.D.I.)

por

Dr. Luis Blanco Felip

Departamento de Pedagogía. Universidad de Lérida

RESUMEN

Una eficaz acción educativa obliga necesariamente a plantearse en profundidad la autoevaluación como modelo de cambio y la participación de todos los estamentos de la comunidad educativa. Un mayor o menor grado de participación viene a ser un indicador que nos permite valorar cualquier propuesta de «calidad», «reforma» o «innovación».

El modelo B.A.D.I. y en especial su instrumento modular, responde a las exigencias particulares de reflexión, análisis y toma de decisiones que se realiza en cada centro para responder adecuadamente a cualquier planteamiento innovador. Los principios fundamentales de este modelo se resumen en:

- a. Es un modelo de concepción organicista, estamental, con definición ideológica.
- b. Responde a un enfoque fundamentalmente rogeriano.
- c. Participa de los enfoques social, abierto, dinámico, cultural y positivo. Comparte la teoría moderna de la organización y sus indicadores se agrupan en cuatro fases: Criterial, presupuestaria, metodológica e informativa.
- d. Se identifica con los modelos centrados en la evaluación de cambio, de forma especial con la autoevaluación.

La flexibilidad que caracteriza al modelo viene dado por el instrumento modular de área o de criterio, como resultado de la participación y el consenso de todos los estamentos de la institución educativa.

ABSTRACT

An efficient educative action necessarily forces an in —depth plan for the self-evaluation as a model for change and the participation of the educative community at all its levels. A higher or lower degree in its participation may be regarded as an indicator that permits us to assess any proposal of «quality», «reform» or «innovation».

The B.A.D.I. model and especially its modular instrument answer to the particular demands of reflexion, analysis and decision-taking that are implemented in each centre to respond adequately to any innovative plan. The principles of the model of reference can be summarized in the following:

- a. It is a model of an organicist conception, sectorial with an ideological definition.
- b. It basically follows the pattern set by Carl Rogers.
- c. It shares a social, open, dynamic, cultural and positive approach. The modern theory of organization is part of it and its indicators are gathered in phases: Criterial, budget-orientated, methodological and informative.
- d. It shows strong links with the models centered in the evaluation of change, especially in self-evaluation.

The flexibility that endows the model can be noted by the area or criteria modular instrument, as a result of the participation and consensus by all the teaching levels of the educational institution.

INTRODUCCIÓN

Cualquier intento evaluador por parte de una institución educativa, requiere un debate sobre los siguientes aspectos:

1. La evaluación de centros como factor de calidad.
2. Variables para un modelo evaluador.
3. Modelos de evaluación de centros.
4. Capacidad de la propia organización para autotransformarse.

1. La evaluación de centros como factor de calidad

Para elaborar la evaluación de un centro educativo, hemos de tener en cuenta que ésta debe fundamentarse en el análisis de cada uno de los elementos que definen el centro. Es decir, la evaluación del profesorado, de sus recursos, la tecnología didáctica, su funcionamiento, equipamiento, nivel sociocultural y económico, carácter público o privado, nº de unidades, etc... La evaluación interdisciplinar de todos y cada uno de los elementos pedagógicos-administrativos configura la evaluación final del rendimiento pedagógico del centro o de la institución, entendida como un conjunto de centros educativos.

Lo que realmente interesa a las instituciones educativas es saber, qué es lo que hace incrementar el rendimiento de sus alumnos. En esta línea los estudios realizados en la década de los setenta centran su atención en la identificación de factores relacionados con la propia organización y funcionamiento de los centros. En la década de los ochenta se constata una gran preocupación por el estudio de la evaluación como herramienta potenciadora de la calidad de la enseñanza, y calidad de los centros educativos. Destacamos a modo de ejemplo el siguiente planteamiento:

«Si los criterios y modos de evaluación que determinan objetivos reales, y en última instancia el producto de la enseñanza y el aprendizaje, son coherentes con los objetivos formalmente establecidos y a través de ellos con los fines generales de la educación y el sistema de valores del que derivan, la evaluación actuará como el más poderoso factor de promoción de la calidad educativa, al garantizar la congruencia y eficacia del sistema en su conjunto» (Orden Hoz, A. de la, 1982).

Para evaluar la calidad de educación de un centro, no podemos centrarnos únicamente en su producto final, es decir en los niveles de rendimiento alcanzados. El control de calidad debe entenderse, como un proceso que acompaña al mismo proceso de educación, procurando prevenir con él, los problemas para que no se produzcan, por lo que cada vez hay que poner mayor empeño en una serie de actividades previas como pueden ser, el diseño del plan de centro, la eliminación de factores perturbadores, estudio previo de los recursos planteados, proyecto de autoevaluación para la innovación del centro educativo, etc...

2. Variables para un modelo evaluador

Dada la complejidad de los centros, muchas y variadas son las variables que pueden ser incluidas en un modelo evaluador. Para Pérez Juste, R. (1988), las variables más importantes son:

- a. Variables referentes al alumnado
- b. Variables referentes al profesorado
- c. Variables del centro escolar
- d. Variables organizativas
- e. Variables relativas al clima de centro

3. Modelos de evaluación de centros

Cuando se trata de evaluar instituciones educativas, los modelos tienden generalmente a asumir las teorías sobre evaluación educativa, así como los esquemas conceptuales y tecnología conocida en este campo: evaluación input-

output, proceso-producto, sumativa-formativa, externa-interna, cualitativa-cuantitativa, etc... Para Pérez Juste, R. y García Ramos, J. M. (1989) podemos agrupar los distintos modelos de evaluación de centros en grandes bloques según pongamos énfasis en:

- a. Los resultados (outputs-salidas)
- b. La relación entradas-salidas
- c. Los procesos internos de la propia organización
- d. Los aspectos culturales de la organización
- e. Los aspectos técnicos de la organización
- f. La relación organización-factores humanos
- g. Los criterios integradores
- h. La capacidad de la propia organización para autotransformarse

4. Capacidad de la propia organización para autotransformarse

Una buena estrategia de evaluación empieza con una comisión de evaluación institucional reducida, diversa y representativa. Su principal objetivo será el definir qué conviene evaluar, para qué y cómo, fijar plazos, seleccionar datos, elaborarlos y extraer consecuencias, finalmente divulgarlos a todos los implicados. Los responsables de la gestión pueden o no implicarse en esta comisión. Algunos principios orientadores para este trabajo pueden ser:

- a. Ajustar el plan de evaluación al plan de la institución
- b. La sensatez y el sentido común práctico han de orientar el trabajo de la comisión
- c. La evaluación de la institución ha de ser a «su medida» y responder a «sus necesidades y posibilidades»
- d. La evaluación de aspectos parciales ha de contextualizarse en el todo integral de la institución
- e. Aprovechar los datos disponibles y analizarlos a fondo
- f. De poco y bien analizados se obtiene más que de mucho y superficial

Nuestro modelo de evaluación, previo estudio de la realidad educativa de nuestro país y dada la necesidad que tendrán las instituciones educativas de adaptarse a las nuevas exigencias sociales, legislativas, económicas, demográficas, hace opción por un modelo autoevaluador, es decir que sea el propio centro escolar (o institución) el que se constituya como evaluador de su funcionamiento y rendimiento a través de los distintos equipos, estructuras, órganos y consejos.

En segundo lugar, la compleja cantidad de elementos que intervienen en las instituciones nos obliga a delimitar las áreas que deben ser evaluadas. Nuestra propuesta presenta un modelo de evaluación diagnóstica que permite a las instituciones educativas conocer su realidad ante un evento como es la Reforma Educativa. Previo estudio de la LODE y de la LOGSE, deducimos que las áreas más

significativas a considerar para realizar una evaluación diagnóstica y con ello deducir si un centro tiene más o menos condiciones para adaptarse a las exigencias que nos plantea la actual Reforma Educativa, todo parece indicar que son las siguientes:

- a.— Definición ideológica
- b.— Definición pedagógica
- c.— Estructura organizativa
- d.— Coordinación
- e.— Clima de centro
- f.— Innovación pedagógica
- g.— Infraestructura. Servicios. Financiación

PRINCIPIOS CONSTITUYENTES

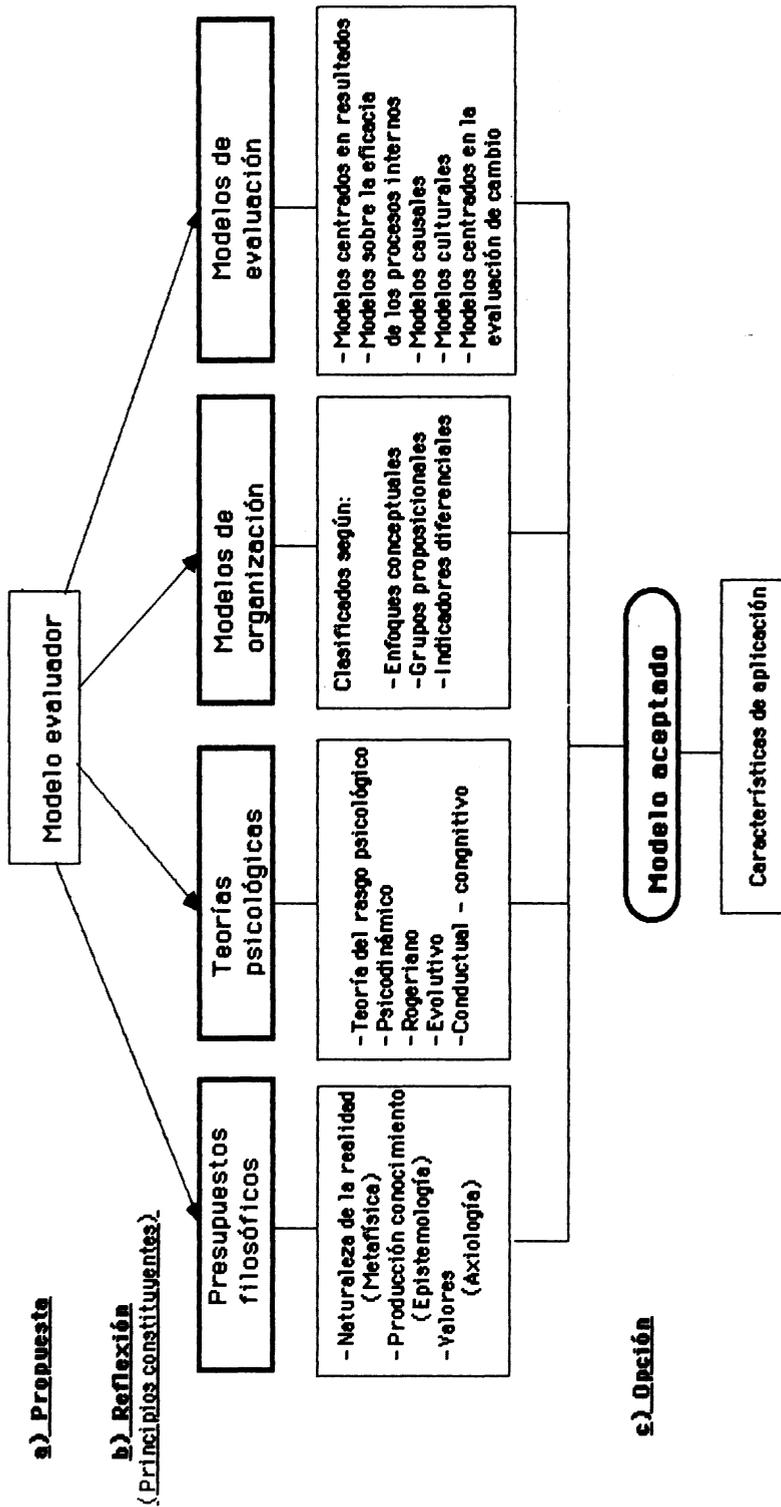
Superada la idea de que el evaluador no es un simple técnico que recoge, analiza y presenta unos resultados, nos encontramos con la realidad de que éstos, objetivamente presentados, de por sí, no tienen ningún valor si no son considerados dentro de su contexto ideológico. Este contexto recupera un protagonismo esencial en el momento de la toma de decisiones. Para la elaboración de un modelo evaluador de un centro educativo debemos considerar los siguientes momentos:

1º. Propuesta: Se plantea la necesidad de hacer un modelo evaluador.

2º. Reflexión: Se realiza el estudio de aquellos principios que pueden configurar el nuevo modelo. Entendemos como principios constituyentes:

- a. Presupuestos filosóficos
- b. Teorías psicológicas
- c. Modelos de organización
- d. Modelos de evaluación

3º. Opción: Previa propuesta y posterior consideración de los principios constituyentes, se establece la opción definitiva del modelo. Éste intentará responder a las expectativas por las que inicialmente fue planteado. La relación que se establece entre los tres momentos mencionados es la siguiente:



1. Principios que constituyen la elaboración de un nuevo modelo evaluador de centros educativos. Propuesta del modelo B.A.D.I.

Realizado el estudio de los principios constituyentes presentados proponemos un modelo de autoevaluación de centros con las siguientes características:

a. Presupuestos filosóficos

* Metafísica

Es un modelo esencialmente organicista, ya que la escuela es una realidad en continua construcción y precisa de una síntesis dialéctica por parte de todos los estamentos de la institución. Conviene complementar dicha concepción con el contextualismo (valor de la realidad histórica del centro), el mecanicismo (la institución constituye un entramado de leyes internas) y el formismo (donde se enfatizan las semejanzas y las diferencias de la organización).

* Epistemología

El conocimiento producido por el nuevo modelo recoge los diferentes niveles de conocimiento, es decir, la colaboración de los distintos estamentos que componen la comunidad educativa. Es pues un modelo estamental.

* Axiología

El modelo a proponer tendrá en consideración la axiología del centro. Es pues un modelo con definición ideológica, cuya referencia estará en el Carácter Propio del centro.

b. Teorías psicológicas

* Enfoque

El nuevo modelo tendrá un enfoque fundamentalmente rogeriano, ya que la relación existente entre la aplicación del modelo y el centro debe ser cordial y participativa, sin ánimos de confrontación. Se pretende una actitud optimista que predisponga al cambio constructivo. A la vez, consideraremos las aportaciones que otros enfoques nos puedan plantear.

* Otros enfoques

a. Teoría de rasgos psicológico

* Adecuación del modelo a las peculiaridades del centro

* Los evaluadores con cierto nivel

b. Psicodinámico

* El nuevo modelo surge de las necesidades de orientación y perfeccionamiento que tiene el centro. Lograr un autoconcepto y una autonomía

c. Evolutivo

* Contemplar el proceso evolutivo de la realidad escolar

d. Conductual-cognitivo

* Interrelación entre factores internos y externos del centro

* Necesidad de una ayuda especializada. El centro vive un determinado período en un determinado contexto

c. Modelos de organización

En cuanto a los esquemas conceptuales sobre organización, nuestro modelo participa fundamentalmente de los siguientes enfoques:

a. Enfoque social

* La estimación del producto va unida a la evaluación de los procesos internos (clima, satisfacción de la organización, mejor calidad de vida en el trabajo...)

b. Enfoque abierto

* Las organizaciones no pueden quedar al margen de sus interacciones con su contexto

c. Enfoque dinámico

* La capacidad de la organización en autotransformarse (posibilitar el desarrollo de la propia organización)

d. Enfoque cultural

* Cada organización tiene su propia cultura. Esta cultura es la que aglutina, aporta energías y crea un clima a la organización

e. Enfoque positivista

* Utilización de un análisis preferentemente cuantitativo para el estudio de la organización escolar

d. Modelos de evaluación

Nuestro modelo se define dentro de los considerados «modelos centrados en la evaluación de cambio». La realidad en que viven las instituciones escolares plantean la inmediatez del cambio. Por ejemplo: ¿tiene la escuela condiciones para una adecuada incorporación a la Reforma Educativa?

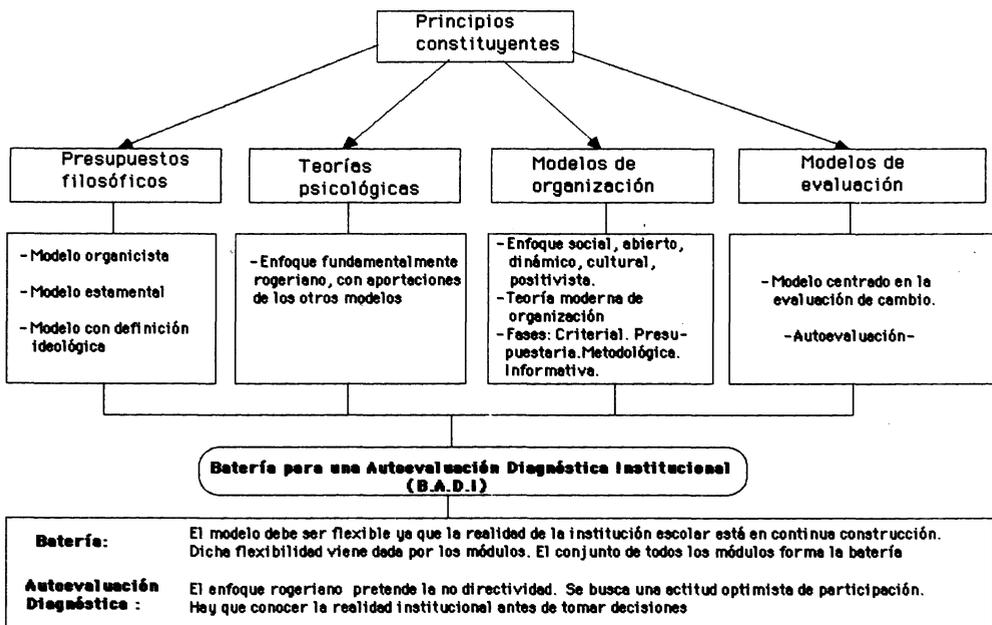
2. Denominación y resumen

Los principios anteriormente señalados se interrelacionan para dar lugar a la formación de un nuevo modelo evaluador de centros educativos que se denominará

«Batería para una autoevaluación diagnóstica institucional (B.A.D.I.)». Los principios que constituyen dicho modelo, se resumen en los siguientes puntos:

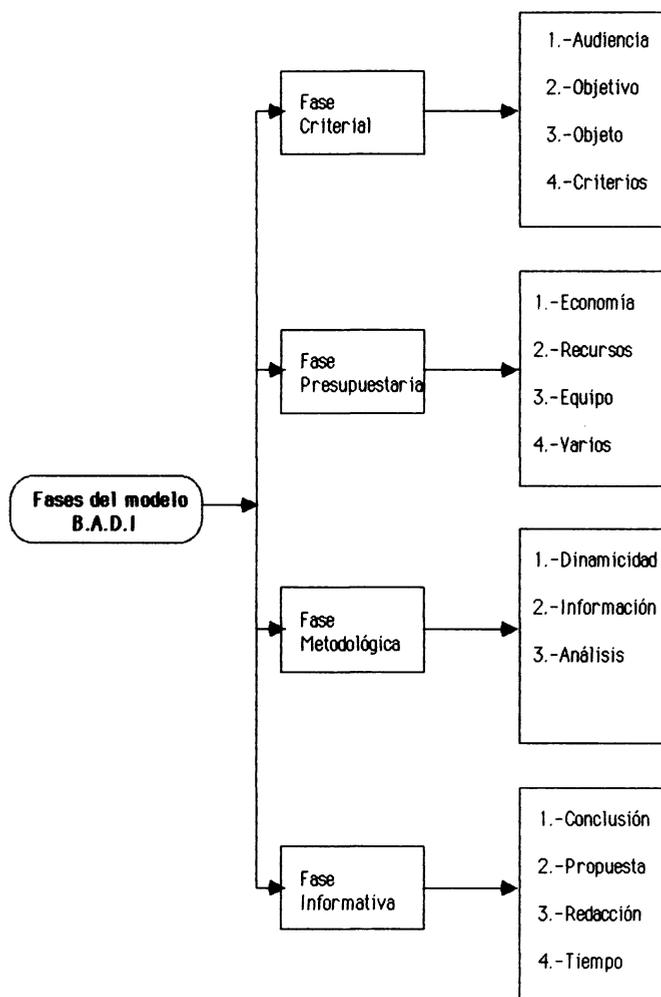
- a. Es un modelo de concepción esencialmente organicista, estamental, con definición ideológica
- b. Tiene un enfoque fundamentalmente rogeriano
- c. Participa de los enfoques conceptuales: social, abierto, dinámico, cultural y positivo. Comparte la teoría moderna de la organización y sus indicadores pueden agruparse en cuatro fases: Criterial. Presupuestaria. metodológica. Informativa.
- d. Se identifica con los modelos centrados en la evaluación de cambio. De forma especial con la autoevaluación

Gráficamente diremos que:



3. Fases de la Batería para una Autoevaluación Diagnóstica Institucional

Los indicadores que componen cada una de las fases del modelo B.A.D.I. se indican a continuación



INSTRUMENTO MODULAR

Entendemos por módulo «al conjunto de elementos que tienen entidad propia para ser evaluados». Cada módulo es independiente entre sí y puede plantearse desde dos aspectos: el instrumento modular de área (módulo de área) y el instrumento de criterio (módulo de criterio). El instrumento modular del modelo B.A.D.I. ofrece una doble posibilidad cualitativa y cuantitativa.

1. Instrumento modular de área

Podemos hablar del instrumento modular de área cuando el objetivo general de evaluación es una o más áreas de identificación. Los pasos a considerar para la aplicación del mencionado instrumento son los siguientes:

- a. Concretar el área de identificación que pensamos evaluar. Ejemplo:
 - * Tutoría
- b. Especificar los indicadores que caracterizan el área seleccionada. Ejemplo:
 - * Relación profesor-alumno
 - * Tiempo de dedicación
 - * Tutoría personal
 - * Tutoría grupal
 - * Programación de la acción tutorial
- c. Determinar el criterio o criterios que pensamos utilizar para la elaboración de los ítems cualitativos y cuantitativos. Ejemplo:
 - * Sinceridad
 - * Utilidad
 - * Confianza
 - * ...
- d. Seleccionar las fuentes de información. Ejemplo:
 - * Alumnos del curso anterior
 - * Alumnos del presente curso
 - * Padres
 - * Profesores no tutores

2. Instrumento modular de criterio

Podemos hablar del instrumento modular de criterio cuando nuestro objetivo es la evaluación de un criterio aplicado a una o más áreas de identificación. Los pasos a seguir para su aplicación son:

- a. Concretar el criterio que pensamos evaluar. Ejemplo:
 - * Utilidad
- b. Especificar las áreas de identificación. Ejemplo:
 - * Escuela de Padres
 - * Seminarios
 - * Actividades complementarias
 - * Tutoría grupal
- c. Determinar los indicadores de cada área. Ejemplo:
 - * Escuela de Padres
 - * Reuniones semanales
 - * Conferencias
 - * Mesas redondas

- * Seminarios
 - * Recursos didácticos
 - * Reuniones
- * Actividades complementarias
 - * Semana Cultural
 - * Festivales
- * Tutoría grupal
 - * Participación
 - * Dinámica
 - * Actualidad de los temas

d. Seleccionar las fuentes de información. Ejemplo:

- * Alumnos de COU
- * Profesores BUP-COU
- * Padres de COU

Resumiendo diremos:

Módulo de área	Módulo de criterio
1. Concretar las áreas de identificación	1. Concretar el criterio o criterios
2. Especificar indicadores	2. Especificar áreas de identificación
3. Determinar el criterio o criterios para elaborar los ítems	3. Determinar los indicadores de cada Área de Identificación
4. Seleccionar fuentes de información	4. Seleccionar fuentes de información

3. Aplicación del instrumento modular del modelo B.A.D.I.

La aplicación metodológica del modelo B.A.D.I. puede optar indistintamente por una evaluación de módulo de área, de criterio o de las dos a la vez ya que no son excluyentes entre sí. Nuestra propuesta se dirige principalmente a una evaluación modular de área, ya que a partir de ésta se puede obtener fácilmente los resultados correspondientes a una evaluación modular de criterio. En la evaluación que se realizó a varios centros educativos, se tomó la L.O.G.S.E. como marco referencial, y se desarrollaron los pasos propuestos en el módulo de área.

1. Módulos de área

- a. Definición ideológica
- b. Definición pedagógica
- c. Estructura organizativa
- d. Coordinación

- e. Clima de centro
- f. Innovación
- g. Infraestructura. Servicios. Financiación

2. Indicadores

- a. Definición ideológica
 - a.1.— Educación democrática
 - a.2.— Comunidad Educativa
 - a.3.— Carácter Propio del Centro
- b. Definición pedagógica
 - b.1.— Proyecto Educativo
 - b.2.— Programación Anual del Centro
- c. Estructura organizativa
 - c.1.— Manual de Régimen Interno
 - c.2.— Organigrama
 - c.3.— Órganos de gobierno unipersonales
 - c.4.— Órganos de gobierno colegiados
- d. Coordinación
 - d.1.— Ciclos y niveles
 - d.2.— Equipos docentes
 - d.3.— Orientación
- e. Clima de Centro
 - e.1.— Participación
 - e.2.— Condiciones laborales
 - e.3.— Condiciones económicas
 - e.4.— Comisión de conflictos
- f. Innovación pedagógica
 - f.1.— Formación permanente
 - f.2.— Experimentación (actividades)
 - f.3.— Evaluación
- g. Infraestructura. Servicios. Financiación
 - g.1.— Edificio
 - g.2.— Patios
 - g.3.— Bus
 - g.4.— Comedor
 - g.5.— Financiación

3. Criterios

- a. Presencia
- b. Utilidad
- c. Opinión

- d. Intervención
- e. Evaluación

4. Fuentes de información

- a. Estamento de Dirección
- b. Estamento de Profesores
- c. Estamento A.P.A.
- d. Estamento P.A.S.
- e. Estamento Alumnos

A modo de ejemplo presentamos los elementos que componen uno de los módulos de área:

- 1. MÓDULO DE ÁREA: Estructura Organizativa
- 2. INDICADORES:
 - * Manual de Régimen Interno
 - * Organigrama
 - * Órganos de gobierno unipersonal
 - * Órganos de gobierno colegiados
- 3. CRITERIOS DE ÁREA:
 - * Presencia e intervención (paradigma cualitativo)
 - * Utilidad, opinión y evaluación (paradigma cuantitativo)
- 4. FUENTES DE INFORMACIÓN:
 - * Estamentos de Dirección, Profesores, A.P.A., P.A.S., Alumnos

Por lo que a este ejemplo se refiere, el número de respuestas que se obtuvieron en un centro educativo fue de 544 respuestas.

MÓDULO: Estructura Organizativa
 INDICADORES: 4
 RESPUESTAS: 544

E.D.	Prof.	APA	PAS	Alumnos
(23x4)	(47x4)	(17x4)	(11x4)	(38x4)
92	+ 188	+ 68	+ 44	+ 152

APLICACIÓN, RESULTADOS Y TOMA DE DECISIONES

Las características del modelo para una autoevaluación institucional permiten adaptarse y responder a las diferentes necesidades planteadas en el campo de la evaluación institucional educativa. Dicha teoría fue confirmada en el ámbito de dos tipos de instituciones educativas. Los resultados de la aplicación del modelo B.A.D.I.

a tres instituciones de educación secundaria (BUP y COU) y a una institución de educación universitaria, fueron totalmente satisfactorios. Las características de adaptabilidad y flexibilidad que tiene el modelo y su instrumento modular fueron suficientemente demostradas.

Después de una aplicación cero, el instrumento modular fue aplicado a una muestra de tres centros de BUP-COU y a un centro de educación universitaria, el cual recogió un total de 19.609 respuestas, que se distribuyen de la siguiente forma

Aplicación cero:	409	respuestas
Aplicación 1, 2 y 3:	13.200	respuestas
Aplicación universitaria:	6.000	respuestas
	19.609	respuestas

1. Sobre la validez del instrumento

Con respecto a la validez del instrumento tomamos como referencia el trabajo realizado por Gronlund, N.E. (1985), «Measurement and evaluation in teaching», destacando los siguientes aspectos.

- a. La validez está relacionada con el uso que se hace de los resultados y la firmeza de sus interpretaciones. Es más correcto hablar de la «validez de interpretación que se hace de los resultados» que hablar de la «validez de instrumento».
- b. No podemos dicotomizar la validez con un «sí» o un «no». Debemos entenderlo en términos graduables. Tomamos como referencia de gradación tres niveles: alta, media y baja. La validez de las interpretaciones fue valorada con un grado de *validez alta*, considerada ésta como un concepto unitario.
- c. Nos apoyamos, según el planeamiento teórico, en tres evidencias:

a. Evidencia de contenido:

1. Con respecto a las áreas propuestas, ninguno de los expertos consultados ha presentado ninguna consideración en contra, por lo que debemos entender que lo que medimos responde a lo que queremos medir.
2. El texto legal de la LOGSE, no contradice la importancia, de los módulos evaluados.
3. La autoevaluación realizada a las aplicaciones A.1, A.2 y A.3 presentaron un nivel de respuesta afirmativa muy alta. La pregunta fue la siguiente: «Los aspectos que se recogen en este cuestionario y entrevista, pueden considerarse fundamentales teniendo en cuenta la LOGSE»
 - * 97,5% para la institución A.1
 - * 97% para la institución A.2
 - * 95,3% para la institución A.3

4. Entre las observaciones recogidas, no existe ningún tipo de inadecuación entre el contenido de lo que medimos y el contenido de lo que pretendemos medir. El vocabulario y las instrucciones han resultado comprensibles y adecuadas.
5. No solamente existe representatividad, sino que además podemos decir que la representatividad es democrática.

b. Evidencia de criterio

A falta de un «criterio de éxito» por lo que nos permita decir si los centros A.1, A.2 y A.3, obtienen una correlación satisfactoria, con referencia a un hipotético «centro modélico», opinamos que en nuestra investigación no nos interesa tanto utilizar como «evidencia de criterio» otro centro, sino utilizar como criterio el referente «*Muy satisfactoria*» (entendiendo que es esta la valoración máxima). De acuerdo con este criterio diremos que la validez es plenamente satisfactoria si observamos los resultados obtenidos:

- A.1: De Satisfactoria a Muy Satisfactoria (3,6)
- A.2: Muy Satisfactoria (4)
- A.3: Satisfactoria (3,3)

c. Evidencia de interpretación

Cualquier defecto en la construcción de la prueba puede dificultar las interpretaciones que se deducen de sus resultados. En nuestra investigación no se da ninguno de los factores que podrían justificar una invalidación, a saber:

1. Instrucciones confusas
2. Vocabulario elegido y estructura de las frases demasiado difíciles
3. Nivel de dificultad de las preguntas
4. Ambigüedad
5. Preguntas inapropiadas para los resultados que se están midiendo
6. Pruebas demasiado cortas
7. Orden incorrecto de las preguntas
8. Muestra identificable de las respuestas

2. Requerimientos de la fiabilidad y naturaleza de la decisión

La fiabilidad hace referencia a la consistencia de los resultados de la evaluación, es decir si obtenemos puntuaciones bastante similares administrando la misma prueba al mismo grupo en ocasiones diferentes. Pero esta posibilidad es en nuestro

caso del todo inviable. De acuerdo con Gronlund N.E. (1985), los métodos de correlación para determinar la fiabilidad son:

* Evaluación y reevaluación	Misma prueba dos veces
* Formas equivalentes	Dar dos formas de la prueba
* Evaluación y reevaluación con formas equivalentes	Dar dos formas de la prueba
* Split Half	Dar la prueba una vez
* Kurder-Richardson	Dar la prueba una vez

El grado de fiabilidad que exigimos en nuestras medidas educacionales depende ampliamente de la decisión a tomar. Por ello los requerimos de fiabilidad con respecto a la toma de decisión podrían resumirse de la siguiente forma:

a. Se requiere alta fiabilidad cuando la decisión:

- * Es importante
- * Es final
- * Es irreversible
- * Es inconfirmable
- * Es concerniente a individuos
- * Tiene consecuencias durables

b. Una baja fiabilidad es tolerable cuando la decisión

- * Es de menor importancia
- * Es tomada en etapas tempranas
- * Es reversible
- * Es confirmable con otra información
- * Conciernen a grupos
- * Tiene efectos temporales

Para resumir diremos que los estimados de fiabilidad pueden variar de acuerdo con el largo de la prueba, la amplitud de las puntuaciones en el grupo evaluado, la dificultad de la prueba, la objetividad de la calificación y el método de estimar la fiabilidad. El grado y tipo de fiabilidad ha de ser buscado en una instancia particular, y depende principalmente de la decisión a tomar. Para decisiones tentativas y reversibles, una baja fiabilidad puede ser tolerable, pero para decisiones finales e irreversibles, debemos ser muy exigentes sobre la fiabilidad de nuestras medidas.

3. El coeficiente Alfa de Cronbach

De acuerdo con las características presentadas, podemos observar que ninguno de los métodos nos ofrece las garantías adecuadas. Si bien el método de Kurder y Richardson es el que podría aproximarse más a nuestras necesidades, presenta el

inconveniente de que nuestro instrumento no es de tipo binario. Nuestra opción final se decide por el coeficiente Alfa de Cronbach ya que nuestros ítems admiten una gradación de la respuesta correcta. Para ello tomamos como referente a Cabrera, F. y Espín, J. (1986): «El coeficiente Alfa desarrollado por Cronbach es una generalización de la fórmula KR-20, cuando los ítems no son de tipo binario, es decir, a pruebas cuyos ítems admiten una gradación de la respuesta correcta.». Los coeficientes de fiabilidad obtenidos fueron los siguientes

Aplicación	Coeficiente Alfa
A.1	0,912
A.2	0,950
A.3	0,975
Inst. Universitaria	0,752

4. Toma de decisiones

Las decisiones que deben tomarse una vez finalizada la autoevaluación deberán tener en cuenta el informe elaborado por la comisión de evaluación. Una vez ordenada y clasificada la información recogida por medio del instrumento modular, pasamos a la redacción del informe que se presentará a la audiencia que la haya solicitado. El informe que propone el modelo B.A.D.I. destaca de forma preferente tres tipos de información:

- a. Puntuación global de los módulos evaluados
- b. Perfil y valoración del centro
- c. Perfil y valoración de los módulos

Cada vez más existen centros que además de una evaluación asistemática y permanente, inevitable y necesaria, establecen procesos de evaluación sistemática. Con periodicidad anual, con mecanismos de participación, con criterios y estrategias, con selección y priorización de aspectos a evaluar. Los centros que optan por esta postura entienden la evaluación como un mecanismo importante de gestión y eficacia.

Evaluar es medir procesos y resultados. También es comparar evolución y cambio de la propia institución, es establecer una serie de indicadores objetivos que permitan contrastar resultados. No para competir, sino para contrastar y entender mejor los procesos dinámicos y poder actuar sobre ellos con mayor acierto. En última instancia evaluar es optimizar, mejorar la propia institución.

Observar, reflexionar, comprobar, son eslabones del proceso de evaluación de un centro. Su proceso es complejo, pero conviene desmitificar los tecnocratismos y las dificultades «instrumentales». Las principales dificultades no son técnicas, sino

políticas, de voluntad, de sensatez, de sentido común, de sensibilidad pedagógica, de flexibilidad y de creatividad. Conviene no perder de vista que la mejor evaluación no es la más técnica y precisa, sino la más operativa, es decir, aquella que selecciona y se centra en obtener datos que los elabora y los divulga para ayudar a tomar y mejorar las decisiones.

REFERENCIAS BIBLIOGRÁFICAS

- BARTOLOMÉ, M. y ANGUERA, M. T. (coord.) (1990): *La investigación cooperativa: vía para la innovación en la universidad*. Barcelona: PPU.
- BEER, V.; BLOOMER, A. y CORPORATION, X. (1986): «Levels of evaluation». *Educational Evaluation and Policy Analysis*, Winter, Vol. 8, nº 4.
- BENNIS, W. G.; BENNE, K. D. y CHIN, R. (1969): *The planning of change*. Holt, Rinehart and Winston: New York.
- BLANCO, L. (1992): *Batería para una autoevaluación diagnóstica institucional (B.A.D.I.)*. Tesis Doctoral. Barcelona. Universidad de Barcelona.
- CABRERA, F. y ESPÍN, J. (1986): *Medición y evaluación educativa. Fundamentos teórico-prácticos*. Barcelona: PPU.
- ESCOTET, M. A. (1984): *Técnicas de evaluación institucional en la educación superior*. Madrid: M.E.C.
- ESCUADERO, T. (1989): «Aproximación pragmática a la evaluación de la Universidad», *Revista de Investigación Educativa* nº 13.
- FERNÁNDEZ, M. J. (1986): «Investigación evaluativa de instituciones universitarias», Tesis Doctoral. Madrid. Universidad Complutense.
- FULLÁN, M. (1982): *The meaning of educational change*. Teachers. New York: College Press.
- GRONLUND, N. E. (1985): *Measurement and evaluation in teaching* (5ª ed.) New York: Mac Millan Publ. Company.
- GUBA, E. G. and LINCOLN, Y. S. (1985): *Effective evaluation*. San Francisco. California: Jossey-Bass.
- HUSEN, T. (1988): «Paradigmas de la investigación en Educación: Un informe del estado de la cuestión», *Aspectos metodológicos de la investigación educativa*. II Congreso Mundial Vasco-Narcea.
- HUTCHINGS, MARCHESE (1990): «Evaluation des resultats de programmes d'estudes. L'expérience des Etats Units», *Gestion de l'ensenyament superior*, vol. 12, nº 1.
- MARTÍNEZ ARAGÓN, L. (1989): «El enfoque empresarial (Modelo de Auditoria) aplicado a la evaluación de centros educativos», *Revista de Investigación Educativa* nº 13.
- MARTÍNEZ, A.; BELDA, J.; MAZZOLA, A. y PERELLO, F. (1989): «El proceso de evaluación institucional en la Escuela de Trabajo Social de la Universidad de Valencia», *Revista de Investigación Educativa* nº 13.
- MATEO, J. (1986): *Proyecto docente e investigador de la Cátedra de Medición y Evaluación Educativa*. Inédito. Barcelona. Universidad de Barcelona.
- MATEO, J. (1989): «L'avaluació en l'organització de centres», *III Curs sobre Avaluació Educativa*. Barcelona. Societat Catalana de Pedagogía.
- MIGUEL, M. de (1989): «Modelos de investigación sobre organizaciones educativas», *Revista de Investigación Educativa* nº 13.

- MILES, M. B. (1974): «Las diez características del centro docente sano», *La Educación hoy* nº 5, pp. 197-198.
- MORALES, P. (1988): *Medición de actitudes en psicología y educación: Construcción de escalas y problemas metodológicos*. San Sebastián: Ttartalo.
- MOSLEY, Ch. (1989): «L'avaluació del professorat», *III Curs sobre Avaluació Educativa*. Barcelona: Societat Catalana de Pedagogia.
- MUNICIO, P. (1990): «Diagnosticar en los centros educativos», *Apuntes de Educación* nº 39. Anaya.
- NEVO, D. (1986): «The Conceptualization of Educational Evaluation: An analytical Review of the Literature» en HOUSE, R.: *New Directions in educational evaluation*. The Fatmer Press. Philadelphia. Pp. 15-29.
- ORDEN, A. de la (1982): *La evaluación educativa*. Buenos Aires: Docencia.
- ORDEN, A. de la (1988): «La calidad de los centros educativos, asunto para un Congreso», *Bordón*, vol. 40, 2, pp. 149-161.
- OWEN, J. M. (1989): «Encouraging Action Through New Orientations to Evaluation», *Studies in Educational Evaluation*, vol. 14. Pergamon Press. Israel. Tel-Aviv University.
- PÉREZ JUSTE, R. (1988): «La evaluación de centros», *Apuntes de Educación* nº 31. Anaya.
- PÉREZ JUSTE, R. y MARTÍNEZ ARAGÓN, L. (1989): *Evaluación de centros y calidad educativa*. Madrid: Cincel.
- PITZ y MACKILLIP (1984): *Decision analysis for program evaluators*. Beverly Hills. Sage.
- RENAUD, G. (1989): «L'avaluació formativa i qualificativa en el batxillerat internacional», *III Curs sobre Avaluació Educativa*. Barcelona: Societat Catalana de Pedagogia.
- RINCÓN, D. del y SANTOLARIA, F. (1989): «Evaluación de centros de reforma», *Revista de Investigación Educativa* nº 13.
- RODRÍGUEZ, B. (1990): *Modelo para la evaluación externa de los centros escolares de E.G.B.* Tesis Doctoral. Oviedo. Universidad de Oviedo.
- SABIRÓN, F. (1990): *Evaluación de centros docentes*. (Modelo, aplicaciones y guía). Zaragoza: Central de Ediciones.
- SCRIVEN, M. (1967): «The methodology of Evaluation», *AERA Monograph Series in Curriculum Evaluation*, 1. Rand Mc Nally. Chicago, pp. 39-83.
- VALLES, A. (1989): *Inspección educativa* (Funciones, medios y técnicas de evaluación). Valencia: Promolibro.
- WORTHEM, B. R. and SANDERS, J. R. (1987): *Educational Evaluation*. Lovigman Inc. New York.

LA POSICIÓN DEL PROFESORADO ANTE EL CAMBIO EDUCATIVO. UN ESCALAMIENTO MULTIDIMENSIONAL NO MÉTRICO DE LOS DISCURSOS SOBRE LA REFORMA

por

Javier Gil Flores

Área de Métodos de Investigación y Diagnóstico en
Educación de la Universidad de Sevilla

RESUMEN

Partiendo del decisivo papel que juegan los profesores en la puesta en práctica de cualquier cambio en educación, nos interesamos por el modo en que los profesores de EGB de la provincia de Sevilla se posicionan en relación a la aplicación de la reforma educativa. Las opiniones, sentimientos, actitudes, expectativas, conductas, etc. de diferentes colectivos de profesores quedan reflejados en discursos obtenidos mediante la técnica de los grupos de discusión. La codificación de los discursos abre paso al recuento de frecuencias, el cálculo de distancias entre los discursos y el empleo del escalamiento multidimensional no métrico. El impacto de la reforma sobre cada sector del profesorado, es decir, las consecuencias que para ellos se derivan de su aplicación, constituye la dimensión respecto a la cual se ordenan los discursos analizados. El artículo finaliza con una valoración de la aplicación del escalamiento multidimensional a datos cualitativos como los manejados en este caso.

ABSTRACT

Taking into account the crucial role that teachers play in the implementation of any educational change, we are interested in the way in which Primary School teachers from the province of Seville think about the application of the educational reform. The opinions, feelings, attitudes, expectations, behaviours, etc. of different groups

of teachers are expressed in the discourses obtained through the focus group technique. The discourses codification let us count the frequencies, calculate the distance between the discourses, and use the non-metric multidimensional scaling. The impact of this reform on each sector of teachers, that is, the consequences that they can derive from its application constitute the respective dimension to which the analyzed discourses are ordered. This paper concludes with an assesment of the application of the multidimensional scaling technique to qualitative data as the ones used in this case.

1. INTRODUCCIÓN

El cambio de la educación resulta necesario en un contexto social en transformación continua. Las nuevas demandas surgidas de una sociedad constitucionalmente democrática, culturalmente plural e integrada en la comunidad internacional, junto con las carencias que venían detectándose en la educación española, han llevado a poner en marcha el actual proceso de reforma derivado de la aplicación de la LOGSE. De acuerdo con la delimitación conceptual establecida por la UNESCO (1981), el cambio educativo adoptaría la denominación de reforma cuando opera al mismo tiempo sobre objetivos, estrategias y prioridades, implicando generalmente un importante cambio estructural. La reforma, en la medida en que representa un cambio educativo a escala de todo el sistema, constituye un proceso político-institucional, y suele concebirse, formularse y planificarse a nivel central (Sack, 1981).

Sin embargo, como afirma McCaig (1981), las reformas implican decisiones políticas, estratégicas que requieren de otras personas distintas a las que las formularon para ser llevadas a la práctica. Parece evidente que la puesta en marcha de los cambios educativos no depende sólo de la regulación legal de los mismos, sino que su concreción práctica se verá mediatizada por el modo en que los profesores asuman y actúen en relación al cambio. Una buena parte de los estudios sobre este tema, se ha dirigido a identificar aspectos que influirían en el éxito de las propuestas innovadoras, tales como las características del profesor, su formación, su conocimiento de las innovaciones, su participación en las decisiones, su coordinación con otros profesores, etc. (Fullan, 1982; Sarason, 1982). Estas concepciones entienden el cambio como un producto acabado y consideran al profesor un aplicador del cambio educativo diseñado por los técnicos; lo importante son las características externas del profesor que podrían estar relacionadas con la adecuada ejecución. Tales planteamientos ignoran las concepciones y las experiencias previas del profesor al enfrentarse a los cambios, cuya consideración llevaría a dirigir la atención hacia el modo en que los profesores asimilan, valoran y transforman las propuestas de cambio (González González, 1984). Para los autores que confieren importancia al punto de vista de los agentes encargados de aplicarlo, resulta prioritario conocer, valorar e interpretar las percepciones del cambio que tienen los profesores, a fin de

poner en marcha acciones facilitadoras del desarrollo de la innovación. Un conocimiento de la posición de los docentes en relación a la reforma educativa constituye el punto de partida para lo que Escudero (1990) denomina «preparar el terreno», y que consistiría en generar un proceso y crear las condiciones para que las propuestas de cambio puedan traducirse efectivamente en la práctica.

Es claro, por tanto, que más allá del esfuerzo planificador de la Administración Educativa, el éxito o fracaso de toda innovación depende en buena medida de los profesores encargados de implementarla. En el caso de nuestra reforma, la adaptación a una nueva estructura del sistema y a una nueva función docente implican una transformación de la actuación profesional del profesorado y de su ubicación como docentes dentro del sistema educativo. Básicamente, la información demandada, las inquietudes e interrogantes del profesorado se han dirigido hacia el modo en que la reforma les afecta, especialmente los aspectos relacionados con la nueva estructuración, nuevos centros, etc.; aspectos sobre los cuales la Administración Educativa ha tratado de tranquilizar a los profesores (Marchesi, 1990). Posiblemente, el mayor reto lo represente la adaptación a un nuevo perfil profesional que la reforma exige del actual profesor. De acuerdo con el papel destacado que se confiere a la profesionalización del docente en las reformas actuales (Popkewitz, 1990), el profesor debe saber reflexionar sobre su práctica; su perfil es el de un profesional capaz de analizar el contexto en que desarrolla su trabajo y diseñar con autonomía las concreciones curriculares adaptadas al centro y a los alumnos (MEC, 1989). En la medida en que el cambio supone asumir un papel más activo en el diseño y desarrollo del currículum, es necesario que se dé en los profesores un proceso de rearme profesional o «reprofesionalización» (Gimeno, 1990). En estas coordenadas habría que entender la necesidad de transformación que el cambio educativo exige de los docentes.

La posición de los profesores ante la reforma estará en buena parte influida por el modo en que, de acuerdo con sus características personales y profesionales, se enfrentan a los elementos que configuran el nuevo marco para su práctica docente, lo cual nos lleva a suponer la existencia de posicionamientos diferenciados en relación a este tema. El trabajo que aquí presentamos se dirige precisamente al objetivo de caracterizar los puntos de vista que mantienen distintos colectivos de profesores a partir de los discursos elaborados en grupos de discusión constituidos por sujetos pertenecientes a tales colectivos.

2. PROPÓSITO DEL ESTUDIO

Tanto el debate inicial en torno al Proyecto para la Reforma de la Enseñanza (MEC, 1987), como la posterior aprobación de la LOGSE y la publicación de propuestas curriculares concretas han hecho surgir un amplio abanico de opiniones, posicionamientos, rechazos o adhesiones de personalidades, colectivos profesionales u organizaciones relacionadas con la escuela. Circunscribiéndonos al ámbito del

profesorado, las medidas administrativas puestas en marcha, tales como la habilitación y adscripción del profesorado o la definición de plantillas, los programas de formación y las ofertas editoriales se han encargado de que difícilmente el tema de la reforma no constituya un asunto por el que de alguna manera se sientan interesados todos y cada uno de los efectivos docentes del sistema escolar.

Pero, como ya hemos avanzado, consideramos que el modo de opinar y responder ante la reforma educativa no es único. De ahí, que nuestro interés se haya centrado en explorar los distintos posicionamientos del profesorado de los centros públicos de EGB y Preescolar, en la provincia de Sevilla, en relación a la implantación de la reforma representada por la LOGSE.

3. PRODUCCIÓN DE DISCURSOS ACERCA DE LA REFORMA EDUCATIVA

La posición del profesorado ante la reforma pensamos que se traduce, en parte, en un conjunto de opiniones, sentimientos, actitudes, expectativas, conductas, etc. a los que podríamos acceder a través de una técnica como la de los grupos de discusión. Los grupos de discusión, ampliamente utilizados en el ámbito de la investigación de mercados, comienzan a estar presentes también en la investigación educativa. Constituyen una técnica de recogida de datos de naturaleza cualitativa, que reúne a un número limitado de personas (entre 7 y 10) desconocidas entre sí y con características homogéneas en relación al tema investigado, para mantener una discusión guiada en un clima permisivo, no directivo (Krueger, 1991). El encontrarse entre personas de características similares y el clima abierto y relajado de la discusión crean el medio adecuado para que se dé un reforzamiento mutuo y lleguen a manifestarse ideas y opiniones, sobre el tema propuesto, que en contextos más rígidamente estructurados o ante personas conocidas no serían manifestadas. La idea básica en los diseños de investigación basados en la técnica de los grupos de discusión consiste en considerar tantos grupos como segmentos de la población, de modo que podamos acceder a los distintos discursos sobre el tema investigado.

Teniendo en cuenta estas consideraciones, nuestro objetivo ha sido producir los discursos de grupo característicos de cada uno de los modos de percibir y entender la reforma educativa. Para determinar el número de colectivos o grupos diferenciados de opinión, y por ende, el de grupos de discusión, se realizó un estudio previo de la población de profesores de EGB de la provincia de Sevilla. En ese estudio, se consideraban los datos de los 7.723 profesores de centros públicos de la provincia respecto a las variables edad, sexo, experiencia docente, titulación, nivel en que imparten docencia e implicación en actividades de innovación (medida a través de la participación en actividades de innovación y experimentación educativas convocadas por la Consejería de Educación de la Junta de Andalucía para el curso 1991/92). Estas variables se asocian a posibles diferenciaciones en las estructuras de opinión y, en particular, a las perspectivas del profesorado ante la reforma: la

influencia de la edad en las necesidades, conductas, actitudes del profesor, percepciones, expectativas, etc. ha sido destacada por autores como Oja (1989) y Sikes (1992); distintos tipos de aspiraciones y expectativas profesionales parecen caracterizar diferenciadamente a profesores y profesoras (Ortega y Velasco, 1991); los estudios sobre problemas y satisfacciones de los profesores revelan que los maestros más expertos mencionan diferentes clases de problemas y satisfacciones que los que cuentan con una experiencia limitada (Hall y Hord, 1987); la titulación y el nivel en que ejercen sitúa en distintas posiciones a los profesores de cara a su promoción o los colocan más o menos próximos en el tiempo a la aplicación de la reforma; por último, la implicación en los cambios ha sido incluida entre los factores para el éxito de cualquier innovación (González González, 1984) y determinaría actitudes de aceptación o rechazo (Esteve Zarazaga, 1991).

A partir de los datos sobre estas variables, proporcionados por el Centro de Proceso de Datos de la Delegación Provincial de Educación, se llevó a cabo un análisis factorial de correspondencias, seguido de una clasificación jerárquica ascendente de las puntuaciones factoriales obtenidas para cada sujeto en los cinco primeros ejes factoriales. El resultado fue la identificación de seis clases o colectivos de profesores que podrían considerar de modo diferente el tema de la reforma educativa. Para establecer las diferencias entre las clases, éstas fueron caracterizadas por ciertas modalidades, analizando la distancia entre el porcentaje de sujetos que presentan determinados atributos dentro de la clase y el porcentaje dentro de la población global.

Un total de 6 grupos de discusión fueron constituidos en otras tantas localidades de la provincia, reuniendo cada uno a profesores seleccionados aleatoriamente entre los que contaban con las características correspondientes a una de las clases. Los seis grupos constituidos se caracterizaban por los rasgos mostrados en el Cuadro 1. Como se refleja en este cuadro, a cada uno de ellos hemos asignado una denominación que refleja alguna de las características más salientes por las que queda definida su composición. Las reuniones se celebraron, en los seis casos, a finales del segundo trimestre del curso escolar 1991/92, con una duración de hora y media aproximadamente. Durante estas reuniones, el moderador se limitó a plantear el tema de discusión y a catalizar la producción del discurso deshaciendo bloqueos y controlando su desarrollo para que se mantuviera dentro del tema. Los discursos elaborados en los grupos fueron registrados mediante una grabadora y transcritos, resultando seis textos para analizar.

CUADRO 1
COMPOSICIÓN DE LOS 6 GRUPOS DE DISCUSIÓN, DEFINIDA A PARTIR DEL ESTUDIO DE LAS MODALIDADES CARACTERÍSTICAS DE CADA CLASE

	COMPOSICIÓN
GRUPO 1 (INNOVADORES)	Profesores con alta implicación en la innovación.
GRUPO 2 (JÓVENES)	Profesores con una edad máxima de 30 años y una experiencia profesional de hasta 2 trienios.
GRUPO 3 (LICENCIADOS)	Profesores destinados en el Ciclo Superior, con una experiencia profesional entre 3 y 7 trienios, varones, en posesión del título de Licenciado y con una edad comprendida entre los 30 y los 45 años.
GRUPO 4 (MEDIOS)	Profesores con una experiencia profesional de 3 a 7 trienios, destinados en los Ciclos Inicial o Medio, con una implicación media en la innovación, con titulación de Diplomados y una edad entre 30 y 45 años.
GRUPO 5 (PROF. PREESCOLAR)	Profesores con experiencia profesional cifrada en 2 o menos trienios, destinados en el nivel Preescolar y con edades desde los 30 a los 45 años.
GRUPO 6 (MAYORES)	Profesores mayores de 45 años, con más de 7 trienios de experiencia profesional y con baja implicación en la innovación.

4. ANÁLISIS DE LOS DATOS

En un primer momento, el análisis de los discursos se llevó a cabo mediante la codificación de los textos, la cual fue realizada de modo abierto e inductivo, por lo que el sistema de categorías no estaba preestablecido sino que surgió como consecuencia del propio proceso de codificación. A partir de una lectura inicial, se identificaron determinados temas, planteándonos en relación a los datos la siguiente

TABLA 1
MATRIZ DE FRECUENCIAS PARA LOS CÓDIGOS EMPLEADOS EN EL
ANÁLISIS

CATEGORÍA	FRECUENCIAS EN LOS GRUPOS							TOTAL
	CÓDIGO	INN	JOV	LIC	MED	PRE	MAY	
Realidad	REA	12	12	5	11	12	9	61
Claridad	CLA	3	1	5	0	2	5	16
Eficacia	EFI	8	16	7	19	2	6	58
Bondad	BON	10	20	17	7	11	9	74
Necesidad	NEC	4	7	2	1	0	5	19
Viabilidad	VIA	7	14	11	7	9	10	58
Escepticismo	ESC	6	5	15	5	7	8	46
Expectación	EXP	0	0	3	1	0	8	12
Esperanza	ESP	0	3	0	1	0	2	6
Temor	TEM	6	4	0	1	6	3	20
Pesimismo	PES	3	2	3	0	0	1	9
Voluntad	VOL	14	3	1	3	6	7	34
Motivación	MOT	15	8	3	5	6	4	41
Información	COG	11	23	7	16	12	9	78
Formación	FOR	9	10	15	15	9	8	66
Participación	PAR	16	13	7	7	17	11	71
Recursos	REC	13	23	9	15	12	5	77
Ratio	RAT	15	8	1	13	10	2	49
Estructura	EST	0	15	2	5	3	8	33
Obligatoriedad	OBL	5	6	8	9	3	1	32
Educación Infantil	EDI	6	3	3	1	26	4	43
Educación Primaria	EDP	3	7	5	2	1	7	25
Educación Secundaria	ESE	5	4	10	1	4	4	28
Currículo	CUR	11	14	18	22	8	13	86
Orientación	ORI	0	0	4	0	1	0	5
Consecuencias alumnos	COA	5	9	1	4	1	0	20
Consecuencias profesores	COP	11	13	11	9	18	8	70
Consecuencias Centros	COC	3	8	0	4	3	3	21
Experimentación	EXR	1	2	0	0	0	4	7
Leyes anteriores	LEA	4	0	8	4	3	7	26
Acuerdo	ACU	28	20	11	17	20	10	106
Desacuerdo	DIS	3	4	13	5	0	5	30
Modelo ideal	IDE	6	6	10	17	3	7	49
Otros temas	ANA	30	52	52	69	52	50	305

cuestión: «¿sobre qué habla este fragmento?». De este modo emerge un conjunto de categorías que es constantemente modificado, redefinido, readaptado en función de los nuevos pasajes del discurso que van siendo objeto de categorización. Los nuevos fragmentos de discurso sirven, en este proceso, para confirmar las categorías existentes o como fuente para la creación de otras nuevas. Un proceso de esta naturaleza exige continuas revisiones del material codificado, comparando las unidades incluidas en cada una de las categorías y comprobando que se da entre ellas una unidad temática. Para facilitar tales tareas, hemos utilizado el programa AQUAD 2.2 (Huber, 1988), diseñado específicamente para el análisis de datos cualitativos, que entre otras funciones, nos permite recuperar todos los fragmentos etiquetados con un mismo código o contar la frecuencia de aparición de los mismos. El sistema de códigos construido y utilizado para la reducción del material cualitativo aparece recogido en la Tabla 1. Las categorías resultantes aluden a valoraciones sobre la reforma, sentimientos y actitudes ante la misma, factores que se consideran clave para su implementación, contenidos estructurales o curriculares, consecuencias derivadas de la aplicación, etc.

Este material codificado permitiría llevar a cabo un análisis de contenido a través del cual acceder a las opiniones vertidas por los profesores en las discusiones de grupo. Sin embargo, nuestro propósito es ahora mostrar el modo en que podría estructurarse la opinión del profesorado de acuerdo con los discursos elaborados sobre la reforma. Las posibilidades que nos abre la cuantificación son interesantes a este respecto.

La naturaleza de los datos cualitativos —palabras, y no números— constituye un problema crónico en la investigación educativa, pues aunque las palabras aportan mayor significado, los números resultan menos ambiguos y son procesados con mayor economía. Por esta razón, es habitual en el análisis del material cualitativo el recurso a la introducción de valores numéricos, que permitan la aplicación de técnicas estadísticas adecuadas a este tipo de datos. Generalmente, los valores cuantitativos surgen a partir del recuento de determinadas unidades diferenciadas en los textos. En este caso, las frecuencias de aparición de los códigos nos han permitido construir la matriz de valores numéricos incluida en la Tabla 1, en la que se recoge el número de veces que aparece cada uno de los códigos en los diferentes discursos que conforman el corpus total de datos.

El propósito de ordenar los diferentes grupos de opinión existentes sobre la reforma, nos ha llevado a situar los discursos de los grupos en un espacio topológico dotado de significado, en el que podamos apreciar las posiciones relativas de unos grupos frente a otros e identificar las dimensiones que sirven para estructurar ese espacio, nos ha llevado a utilizar la técnica del escalamiento multidimensional. El escalamiento multidimensional nos permite visualizar la estructura de un conjunto de objetos al posicionarlos en un espacio de n dimensiones, partiendo de medidas de disimilitud entre los mismos. Dada la naturaleza de nuestros datos, de carácter nominal, hemos optado por el escalamiento multidimensional no métrico, que se apoya en las relaciones ordinales entre los objetos y no en medidas estrictamente

numéricas. De acuerdo con el algoritmo ideado por Kruskal (1964), se parte de una configuración aleatoria de los objetos en un espacio de dimensionalidad fijada, las medidas de disimilaridad se ordenan por rangos, y se transforman en distancias que conservan la misma ordenación que las disimilaridades de partida. La diferencia entre las distancias iniciales y finales nos proporciona una medida de la bondad de ajuste. Esta medida va disminuyendo con sucesivas repeticiones del proceso, hasta conseguirse un valor mínimo, y consiguientemente el máximo ajuste posible del modelo a los datos iniciales.

El escalamiento multidimensional no métrico proporcionado por el programa MDSCAL de Orloci y Kenkel (1987), utilizado en este trabajo, desarrolla el algoritmo de Kruskal (Kruskal, 1964), siguiendo un proceso iterativo que se detiene cuando la disminución del estrés (medida de bondad de ajuste para la representación) conseguida con el último ciclo resulta inferior al 0.001%, o cuando el estrés final correspondiente a la solución es inferior a 0.01, que podría significar una solución total o parcialmente degenerada (Kruskal y Wish, 1986:51). Valores del estrés próximos a cero indican una alteración de la representación debida al efecto del agrupamiento de diferentes puntos.

El estrés constituye un importante elemento no sólo para valorar el grado de ajuste conseguido mediante la representación, sino también para la decisión sobre la dimensionalidad de la misma. Su valor se construye a partir de la suma de diferencias cuadradas entre las distancias en el espacio tras cada iteración (d_{ij}) y las distancias iniciales (D_{ij}). El valor del estrés (S) viene dado por la expresión

$$S = \frac{\sum_{i \neq j}^n (d_{ij} - D_{ij})^2}{\sum_{i \neq j}^n d_{ij}^2}$$

donde n es el número total de objetos sometidos al escalamiento, que en el caso de los discursos sobre la reforma alcanza el valor 6.

Para ser aplicada, la técnica del escalamiento multidimensional precisa de una matriz de distancias entre los objetos que van a ser sometidos al escalamiento. Cuando no contamos con esta matriz, es necesario obtenerla a partir de los datos disponibles acerca de cada uno de los objetos. Hemos utilizado los perfiles de frecuencias de cada uno de los grupos para calcular, de acuerdo con un criterio de distancia determinado, las disimilaridades existentes entre los discursos de los grupos de discusión. En este caso, contamos con la matriz de códigos recogida en la Tabla 1, que ofrece la frecuencia de cada código en los seis discursos considerados. Estos valores han permitido establecer la distancia entre los discursos, de modo que la matriz rectangular de 34 filas por 6 columnas ha sido transformada en otra matriz cuadrada de 6 filas por 6 columnas, recogiendo medidas de distancia entre cada pareja de discursos. Se trata de obtener lo que Dillon y Goldstein (1984) denominan «similitudes derivadas» a partir de la comparación de los perfiles frecuenciales de los discursos.

Habitualmente se emplean diversos procedimientos para obtener la medida de las distancias entre las columnas (o filas) de una tabla. En la literatura se cita la obtención de medidas de proximidad a partir de coeficientes de correlación entre las columnas o a partir de las distancias euclídeas principalmente (Weinberg, 1991). Otros cálculos de disimilitudes entre los perfiles pueden llevarse a cabo de acuerdo con las distancias de Mahalanobis y de Minkowski (Dunn-Rankin, 1983). En el presente estudio, hemos considerado un caso particular de la distancia de Minkowski, cuando el denominador de su exponente adopta el valor unidad.

$$d_{ij} = \{ \sum_{t=1}^r |X_{it} - X_{jt}|^n \}^{1/n}, \quad \text{para } n=1$$

En esa situación, la disimilaridad entre dos columnas podría obtenerse sumando los valores absolutos de las diferencias entre las casillas de cada fila. Es decir, la distancia entre dos columnas i y j vendría dada por la expresión

$$d_{ij} = \sum_{t=1}^r |X_{it} - X_{jt}|$$

donde X_{it} y X_{jt} son los valores presentes en la fila t para las columnas i y j respectivamente, y r es el número de filas de la tabla.

5. RESULTADOS

Los primeros resultados que presentamos corresponden al cálculo de las distancias entre los grupos a partir de los perfiles frecuenciales de cada uno de ellos. La matriz obtenida aplicando el subprograma METRICS, incluido en el programa MDSCAL (Orloci y Kenkel, 1987), se muestra en la Tabla 2. Esta matriz es, por tanto, la utilizada como entrada para el análisis.

TABLA 2
DISTANCIA ENTRE LOS DISCURSOS CALCULADA A PARTIR DE LA
MATRIZ DE FRECUENCIAS DE CÓDIGOS

	DIS.1	DIS.2	DIS.3	DIS.4	DIS.5	DIS.6
DIS.1	0					
DIS.2	162	0				
DIS.3	194	188	0			
DIS.4	175	155	163	0		
DIS.5	124	160	182	171	0	
DIS.6	152	154	118	167	142	0

La decisión sobre el número de dimensiones con que contará el espacio en el que vamos a situar los objetos analizados es uno de los aspectos cruciales en el escalamiento multidimensional. Un modo de establecer la dimensionalidad consiste en observar la evolución del estrés en soluciones con diferente número de dimensiones. Cuando la adición de nuevas dimensiones no consiga una mejora sensible del estrés, y por tanto, de la bondad de ajuste, habremos alcanzado la dimensionalidad más adecuada para representar el conjunto de datos. Gráficamente es posible visualizar esta prueba representando el estrés como una función de la dimensionalidad. Cuando encontremos un «codo» en el gráfico, sospecharemos que una nueva dimensión añade una mejora en el ajuste que puede ser despreciada (Weinberg, 1991:26). En la Figura 1 se pone de manifiesto que, a partir de dos dimensiones, la mejora en la bondad de ajuste se ve considerablemente reducida.

FIGURA 1
RELACIÓN ENTRE NÚMERO DE DIMENSIONES Y ESTRÉS EN EL ESCALAMIENTO MULTIDIMENSIONAL DE LOS DISCURSOS

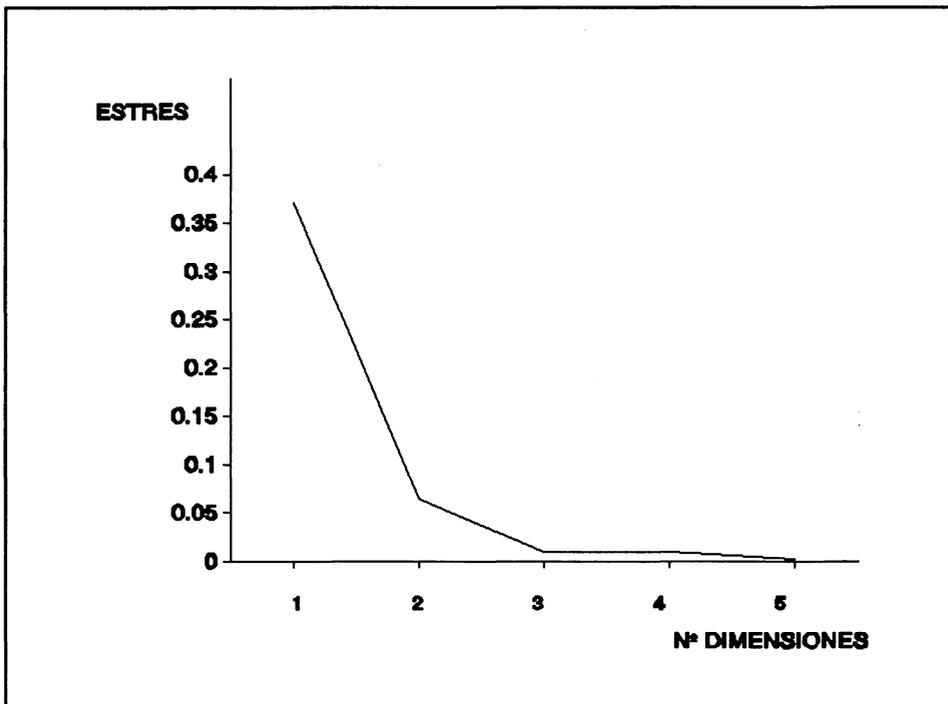
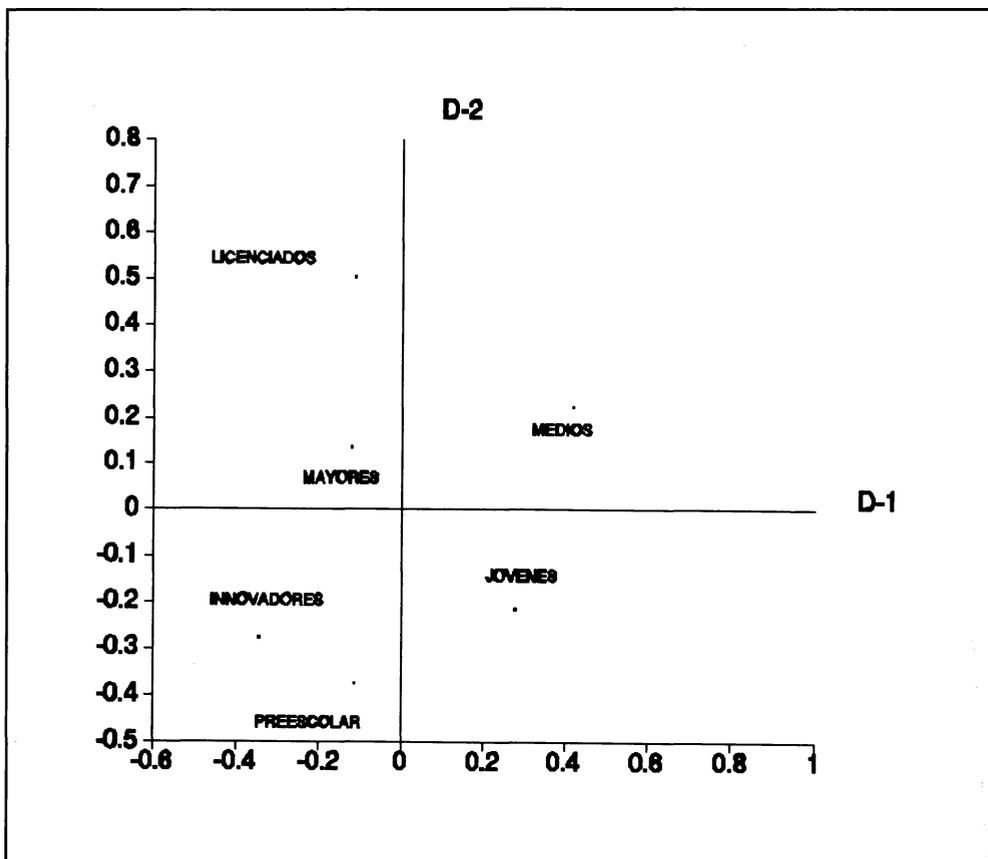


TABLA 3
SOLUCIÓN BIDIMENSIONAL PARA LOS 6 DISCURSOS

GRUPO	EJE 1	EJE 2
GRUPO 1 (INNOVADORES)	-0,3448	-0,2761
GRUPO 2 (JÓVENES)	0,2776	-0,2151
GRUPO 3 (LICENCIADOS)	-0,1158	0,5039
GRUPO 4 (MEDIOS)	0,4197	0,2265
GRUPO 5 (PREESCOLAR)	-0,1147	-0,3746
GRUPO 6 (MAYORES)	-0,1221	0,1355

FIGURA 2
REPRESENTACIÓN GRÁFICA CORRESPONDIENTE A LA SOLUCIÓN BIDIMENSIONAL



Kruskal y Wish (1986:56) estiman que, salvo en el caso de la unidimensionalidad, el estrés no debe situarse por encima de 0.10. Estos autores proponen, además, la interpretabilidad o sentido conceptual de las dimensiones obtenidas como criterio para determinar el número de éstas, y sugieren, en favor de la comodidad de uso, las soluciones de dos o tres dimensiones. Por otra parte, es habitual considerar que el número de dimensiones sea inferior a un tercio de los objetos (Camacho y Hernández, 1986), afirmación que en nuestro caso aconseja descartar una solución tridimensional.

Teniendo en cuenta estos criterios, en el presente estudio hemos fijado un número de dos dimensiones. En la solución obtenida, a la que se llega tras 23 iteraciones, el estrés alcanza un valor de 0.063787 o del 6.38%, indicando el error entre las distancias en la representación y las disimilitudes en la matriz inicial. La posición de los discursos producidos en los grupos de discusión ha quedado determinada por las coordenadas finales resultantes del proceso analítico (ver Tabla 3). La representación gráfica de estas coordenadas origina la Figura 2, donde los discursos quedan posicionados en un espacio de dos dimensiones.

6. INTERPRETACIÓN

En líneas generales, la interpretación de la solución obtenida en un escalamiento multidimensional consiste en identificar agrupamientos en el espacio u ordenamientos a lo largo de una dimensión, para posteriormente describir el rasgo en común de los objetos agrupados y etiquetar el atributo según el cual se ordenan los objetos (Davison, 1983). El primero de los enfoques interpretativos parece no resultar relevante en este caso, ya que contamos con un número reducido de objetos, entre los cuales no parece evidenciarse ningún tipo de agrupamiento espacial. En cuanto al segundo de los enfoques, el modo más simple de interpretar el significado de una dimensión y detectar un posible ordenamiento es intuir las propiedades de los objetos situados en sus extremos, tratando de identificar alguna característica que varíe en sentido opuesto según nos desplazemos desde el origen a uno u otro extremo del eje considerado.

En esta gráfica, los puntos se extienden con mayor amplitud a lo largo del eje de ordenadas, existiendo una mayor concentración de los mismos en lo que se refiere al eje de abscisas. Por tanto, resultará más fácil interpretar la dimensión correspondiente al eje vertical. El grupo de Licenciados se sitúa en el extremo superior de éste, mientras que a Preescolar e Innovadores corresponden las ordenadas más bajas. Basándonos en el modo en que estos grupos son afectados por la aplicación de la reforma educativa, hemos relacionado esta dimensión con el impacto que la reforma tendrá en la situación de los profesores.

Así, el impacto positivo que se atribuiría al grupo de Licenciados viene justificado por las posibilidades de promoción, o al menos de permanencia en los niveles educativos donde se encuentran actualmente. La LOGSE establece, en la disposi-

ción adicional décimosexta, una reserva del 50% de las plazas convocadas para el ingreso en los cuerpos de profesores de Enseñanza Secundaria a los maestros que cuenten con titulación superior, y la disposición transitoria cuarta garantiza la presencia en primer ciclo de la Educación Secundaria Obligatoria a los actuales profesores del ciclo superior. A esto se une la vía que abrirá el desarrollo de la orientación educativa.

El discurso de los profesores en los que la reforma del sistema educativo tendría un impacto positivo, de acuerdo con su posición en el espacio topológico respecto a la dimensión conceptualizada por este rasgo, estaría en consonancia con las posibilidades de promoción que a este profesorado se le ofrecen. Así, como refleja la Tabla 1, es en el discurso de este grupo donde en mayor número de ocasiones se habla de la Educación Secundaria (10 apariciones del código ESE) o de la orientación educativa (4 apariciones del código ORI sobre un total de 5 para todos los grupos).

Menos favorecidos, aunque tampoco perjudicados, se encontrarían los profesores Medios, ubicados en el actual ciclo medio, y que no ven en peligro los puestos docentes que ocupan. La reducción de la Educación Primaria a seis años conlleva una disminución de los cursos impartidos por los profesores de EGB, que a los sujetos de este grupo parece no afectarles directamente. Quedarían también en una zona intermedia, los profesores de elevada experiencia, a quienes la reforma podría ser indiferente, acostumbrados a ver cómo se han desarrollado cambios anteriores o con el horizonte de una pronta jubilación que les excluye de cualquier situación nueva que pueda crearse. Y también aparecen en una zona central los profesores Jóvenes, aún no suficientemente instalados en el sistema como para aferrarse a sus destinos docentes o sentirse perjudicados por su transformación, aunque tampoco sin demasiados motivos para considerarse favorecidos.

Las posiciones en la parte inferior del plano corresponden a los discursos de grupos que podrían recibir un impacto negativo tras la aplicación de la reforma. A los profesores de Preescolar corresponde la ordenada más baja y se encuentran, por tanto, entre los más negativamente afectados. Un importante elemento del cambio a nivel estructural lo constituye la regulación de la Educación Infantil, que queda configurada en dos ciclos. El segundo de los ciclos, el que comprende desde los tres a los seis años, es impartido por maestros, lo cual representa para los profesores de Preescolar el desarrollo de su labor con un alumnado de más corta edad y con necesidades educativas diferentes. Además, se encuentran en los niveles que primero comenzarán a aplicar el cambio. En menor medida, los profesores innovadores quedarían perjudicados, posiblemente al ver institucionalizado un cambio en el que sus expectativas han quedado frustradas por la desconexión entre los esfuerzos de renovación realizados desde la base y los no coincidentes planteamientos que establecerá la Administración.

El discurso, más receloso, de los grupos en los que la reforma tendría un impacto negativo refleja esta circunstancia registrando el mayor número de expresiones de temor: 6 apariciones del código TEM en cada uno de los discursos de los profesores de Preescolar e Innovadores (ver Tabla 1).

7. CONCLUSIONES

La posición de los profesores en relación a la reforma educativa, de acuerdo con los resultados de este estudio, estaría determinada por las repercusiones que la reforma puede suponer para ellos. La preocupación acerca de cómo los cambios les afectan personalmente, aquí evidenciada, es característica de los estadios previos a la aplicación de los cambios (Hall y Hord, 1987). Posiblemente el comienzo de la aplicación lleve a estructurar los discursos del profesorado a partir de otras dimensiones basadas en la experiencia de cambio más que en la anticipación de sus consecuencias. De hecho, la modificación de las opiniones y actitudes es un proceso que tiene lugar en el profesor como consecuencia de las transformaciones observadas en sus prácticas y de los resultados de aprendizaje que obtiene con el uso de la innovación (Guskey, 1986). En cualquier caso, la Administración Educativa tendría que continuar haciendo un esfuerzo por despejar las incógnitas sobre cómo se verán afectados los profesores y clarificar el proceso de aplicación, de modo que la inicial preocupación por las consecuencias del cambio den paso al interés por conocer en profundidad y poner en práctica los contenidos de la reforma.

Algunas consideraciones deben hacerse respecto al modo en que se ha llegado a estas conclusiones. Pensamos que el escalamiento multidimensional no métrico puede ofrecer resultados interesantes al ser aplicado a datos de naturaleza cualitativa como los que hemos manejado en este trabajo. No pueden olvidarse las críticas siempre reiteradas a propósito de la inadecuación de los índices numéricos para reflejar la complejidad de los fenómenos estudiados por las ciencias humanas, o la reducción de información que implican. También se ha señalado a los resultados de un escalamiento multidimensional como el punto de partida para la obtención de conclusiones subjetivas. Incluso Kruskal y Wish (1986) reconocen que los resultados están subordinados a lo que denominan una «interpretación creativa».

Sin embargo, pensamos que trabajar con las distancias entre los discursos, permite al analista comparar y estructurar la información global contenida en el conjunto de datos, tomando los discursos completos como unidad de análisis. Los análisis de datos textuales, como es el caso de los considerados en el presente estudio, que se basan en la codificación del material cualitativo, suelen hacer hincapié en los contenidos intracategorías o en las relaciones intercategorías, pero con mayor dificultad pueden llegar a conclusiones sobre unidades narrativas completas. El escalamiento multidimensional ofrece, desde nuestro punto de vista, una solución satisfactoria en la respuesta a cuestiones como la que en este trabajo nos hemos planteado.

BIBLIOGRAFÍA

- CAMACHO, J. y HERNÁNDEZ, B. (1986): Introducción al escalamiento multidimensional y sus aplicaciones en psicología ambiental, en JIMÉNEZ, F. y ARAGONÉS, J. I. (Comp.): *Introducción a la psicología ambiental*. Madrid, Alianza Editorial, pp. 367-386.
- DAVISON, M. L. (1983): *Multidimensional scaling*. New York, John Wiley & Sons.

- DILLON, W. R. y GOLDSTEIN, M. (1984): *Multivariate analysis*. New York, John Wiley & Sons.
- DUNN-RANKIN, P. (1985): *Scaling methods*. Hillsdale, New Jersey, Lawrence Erlbaum Associates, Publishers.
- ESCUADERO, J. M. (1990): ¿Dispone la reforma de un modelo teórico?, en *Cuadernos de Pedagogía*, 181, 88-92.
- ESTEVE ZARAZAGA, J. M. (1991): Los profesores ante la reforma, en *Cuadernos de Pedagogía*, 190, 54-58.
- FULLAN, M. (1982): *The meaning of educational change*. New York, Teacher College Press.
- GIMENO, J. (1989): Profesionalidad docente, currículum y renovación pedagógica, en *Investigación en la Escuela*, 7, 3-21.
- GONZÁLEZ GONZÁLEZ, T. (1984): Innovación curricular: desarrollo de programas y la figura del profesor, en ESCUDERO, J. M. y GONZÁLEZ, T.: *La renovación pedagógica: algunos modelos teóricos y el papel del profesor*. Madrid, Editorial Escuela Española, pp. 93-164.
- GUSKEY, T. R. (1986): Staff development and the process of teacher change, en *Educational Researcher*, May, 5-12.
- HALL, G. E. y HORD, S. M. (1987): *Change in schools. Facilitating the process*. New York, State University of New York Press.
- HUBER, G. (1988): Análisis de datos cualitativos: la aportación del ordenador, en MARCELO, C. (Coord.): *Avances en el estudio del pensamiento de los profesores*, Sevilla, Servicio de Publicaciones de la Universidad de Sevilla, pp. 77-85.
- KRUEGER, R. A. (1991): *El grupo de discusión. Guía práctica para la investigación aplicada*. Madrid, Pirámide.
- KRUSKAL, J. B. (1964): Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis, en *Psychometrika*, 29, 1-27.
- 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.
- MARCHESI, A. (1990): Profesores, centros docentes y calidad de la educación, en *Cuadernos de Pedagogía*, 184, 10-14.
- MCCAIG, R. (1981): La reforma de la educación: su dimensión humana, en *Perspectivas*, XI, 1, 79-91.
- MINISTERIO DE EDUCACIÓN Y CIENCIA (1987): *Proyecto para la reforma de la enseñanza*. Madrid, Servicio de Publicaciones del MEC.
- MINISTERIO DE EDUCACIÓN Y CIENCIA (1989): *Plan de investigación educativa y de formación del profesorado*. Madrid, Servicio de Publicaciones del MEC.
- OJA, Sh. N. (1989): Teachers: ages and stages of adult development, en HOLLY, M. L. y McLOUGHLIN, C. S. (Eds.): *Perspectives on teacher professional development*. London, The Falmer Press, pp. 119-154.
- ORLOCI, L. y KENKEL, N. C. (1987): *Data analysis in population and community ecology*. Draft copy, University of Hawaii, Honolulu & New Mexico State University, Las Cruces.
- ORTEGA, F. y VELASCO, A. (1991): *La profesión de maestro*. Madrid, CIDE.
- POPKEWITZ, T. S. (1990): Profesionalización y formación del profesorado, en *Cuadernos de Pedagogía*, 184, 105-110.
- SARASON, S. B. (1982): *The culture of the school and the problem of change*. Boston, Allyn and Bacon, Inc.
- SIKES, P. (1992): The life cycle of the teacher, en BALL, S. J. y GOODSON, I. F. (Eds.): *Teachers' lives and careers*. London, The Falmer Press, pp. 27-60.
- UNESCO (1981): *La reforma de la educación: experiencias y perspectivas*. París, Unesco.
- WEINBERG, S. L. (1991): An introduction to multidimensional scaling, en *Measurement and Evaluation in Counseling and Development*, 24, april, pp. 12-36.

TRABAJOS MONOGRÁFICOS

Revista Investigación Educativa - N.º 21 - 1993 (P. 83-106)

LA COMPETENCIA SOCIAL COMO EXIGENCIA ESCOLAR: TÉCNICAS Y PROCEDIMIENTOS PARA SU EVALUACIÓN

por

Gregorio Rodríguez, Eduardo G^a Jiménez y Carmen G^a Pastor
Dpto. de Didáctica y Organización Escolar y Métodos de Investigación
y Diagnóstico en Educación. Universidad de Sevilla

RESUMEN

Centramos este trabajo en la descripción y análisis de las distintas técnicas y procedimientos de evaluación que actualmente se vienen utilizando como medida de la competencia social, aspecto éste que nos servirá de base para plantear posibles vías futuras de investigación.

Descriptores: Competencia social, habilidades sociales, evaluación, evaluación de programas, socialización, interacción social, clima escolar.

ABSTRACT

In this paper we present an analysis and description of different evaluation instruments/strategies that has been used for asses social competence. So we present different ways to further research.

Descriptors: Social competence, social skills, evaluation, programs evaluation, socialization, social interaction, school climate.

INTRODUCCIÓN

El interés por las dimensiones sociales de la enseñanza viene caracterizando la línea de investigaciones que seguimos en el Grupo de Investigación ISIS. Este hecho viene motivado en gran medida porque este grupo ha nacido y se ha desarro-

llado centrado en la investigación sobre integración, ámbito en el que los tópicos ajuste social, habilidad social, aceptación social, se han tomado como referencias del éxito o fracaso en la valoración de este tipo de experiencias. Sin embargo, y a pesar de que la integración sigue siendo la línea de investigación fundamental del grupo, hemos creído necesario trabajar con unas perspectivas más amplias que, a la vez que enriquecieran al grupo, supusieran un mayor énfasis en la unificación de las investigaciones de los campos de educación especial y educación general.

Entre los trabajos realizados en que venimos abordando esta dimensión social podemos citar la investigación titulada «*Adaptación de la enseñanza y cambio de actitudes para la integración de niños deficientes visuales en clases integradas*», dentro de la cual estudiamos la situación de los niños integrados (G. Pastor y G. Jiménez, 1991). También hemos trabajado con el clima de opinión en los centros de integración (G. Pastor y G. Jiménez, 1990), adaptando inventarios de actitudes como el de Larrivee y Cook (1979). Más recientemente, en un nuevo proyecto de investigación, que aún se encuentra en fase de desarrollo, hemos vuelto a trabajar sobre el tema de los alumnos integrados y se está desarrollando un trabajo titulado «*Relaciones psicosociales en las aulas integradas*» (Giraldo, 1992).

Pero es nuestra idea acometer un tipo de investigación que aporte no sólo una comprensión de lo que ocurre con fenómenos educativos como la integración, sino también que de lugar a un determinado tipo de intervención capaz de responder a las necesidades detectadas. En esta nueva línea de trabajo hemos aunado intereses con otros investigadores de nuestra comunidad, preocupados por los problemas de «adaptación social» detectados en muchas de nuestras escuelas, sobre todo en aquellos centros que se han catalogado como «Centros de Actuación Educativa Preferente»¹. Es precisamente desde este ámbito, desde donde se nos hace la propuesta de iniciar un proyecto de investigación que tiene como finalidad la elaboración de programas para el desarrollo de las habilidades sociales (Rodríguez, 1991-92).

Desde nuestra perspectiva hemos tratado de «empezar por el principio», revisando literatura y programas sobre habilidades sociales, aunque a partir de todo ello hemos planteado la necesidad de separarnos de esta línea clásica de intervención, planteando nuestro propio modelo basado en una concepción ecológica de las intervenciones sociales (García Jiménez, García Pastor y Rodríguez Gómez, en prensa).

Uno de los primeros problemas con los que nos encontrábamos es que la dimensión socializadora de la educación se considera implícita en el currículum escolar, pero no se afronta, por sí mismo, el desarrollo de la socialización a través de determinadas estrategias de enseñanza, mediante las cuales los alumnos tengan la oportunidad de aprender diferentes formas de relación social, incluyendo la com-

1 Los Centros de Actuación Educativa Preferente son aquellos que, por las características socio-culturales y económicas del medio en el que se encuentran, son catalogados como tales por parte de la Consejería de Educación y Ciencia de la Junta de Andalucía, con el objetivo de mejorar la calidad educativa de los mismos.

preensión de diferentes perspectivas, el cumplimiento de normas, el desarrollo de roles, etc. Sin embargo, cuando en el medio escolar se procede a la evaluación de los logros de los alumnos, se valoran de forma implícita aspectos que aluden a lo que se considera «competencia social».

El trabajo en equipo, el intercambio de información, completar tareas, responsabilizarse en el trabajo... etc., no son más que aspectos parciales de la competencia social, y su importancia viene avalada por la relación que a lo largo de las múltiples investigaciones realizadas en este campo se han encontrado entre este tipo de competencia y otras características fundamentales para el desarrollo personal y social de los sujetos, tales como el rendimiento escolar, las relaciones interpersonales y la adaptación social (Coie y Krehbiel, 1987; Parker y Asher, 1987; Elliot, Sheridan y Gresham, 1989; Frenzt, Gresham y Elliot, 1991). Así pues, la evaluación de este nivel de competencia se hace necesario en cuanto:

- a) La identificación de los problemas y dificultades en el desarrollo de la competencia social nos permite conocer el comportamiento social de los niños, pudiendo identificarse, de esta forma, las necesidades a que ha de responder la intervención educativa.
- b) El proceso de evaluación nos permitirá, también, la identificación de las áreas, contextos y sujetos sobre los que será necesario incidir de una forma más específica.
- c) La eficacia de un programa de intervención educativa vendrá determinada por el grado en que la misma sea capaz de mejorar las áreas deficitarias que se hayan identificado al iniciar el programa.

Este trabajo lo centraremos en la descripción y análisis de las distintas técnicas y procedimientos de evaluación que actualmente se vienen utilizando como medida de la competencia social, aspecto éste que nos servirá de base para plantear posibles vías futuras de investigación.

1. CUESTIONES PREVIAS

Como paso previo al análisis de instrumentos/procedimientos y estrategias de evaluación consideramos fundamental tratar algunas cuestiones previas que, en el caso de la evaluación de la competencia social, requieren una atención especial, como son la *validez social*, *las modalidades de evaluación*, *la practicabilidad*, *la información que ofrece* y *las cualidades técnicas (psicométricas)*, así como los *marcos teóricos* que subyacen al concepto de competencia social.

a) Validez Social

En primer lugar, debemos considerar que, al venir trabajando en el campo de la competencia social, la evaluación debería reflejar y ser de interés al medio social del niño. En consecuencia, tanto en la evaluación como en la enseñanza, han de con-

templarse aquellos comportamientos que son considerados significativos e importantes por parte del medio social en el que se desarrolla el niño; en definitiva, la evaluación ha de poseer *validez social* (Gresham, 1986; Michelson, 1987).

«El hecho de seleccionar un comportamiento que puede que no sea importante para el medio ambiente del niño (por ejemplo, dar la mano para saludar a los demás) supondría centrar la evaluación y la enseñanza en habilidades sociales cuya relevancia es mínima, o ninguna, para las interacciones sociales reales del niño» (Michelson, 1987:30).

b) Modalidades de evaluación

Como segundo aspecto a contemplar reseñábamos anteriormente las modalidades de evaluación, tema sobre el que nos centraremos con una mayor profundidad en el apartado dedicado a las técnicas y procedimientos de evaluación, pero sobre el que, en principio, debemos tener presente que no existe una única modalidad de evaluación. Así, podremos contemplar desde las técnicas sociométricas hasta la observación natural, dependiendo la elección de unas u otras de los objetivos y posibilidades de la propia evaluación.

«No existe una batería de tests o métodos estandarizados para evaluar las habilidades sociales. Más bien, las hipótesis generadas orientan la dirección de la evaluación, las preguntas que han de ser respondidas, y los métodos que deben utilizarse» (Elliot, Sheridan, y Gresham, 1989:200).

c) Practicabilidad

Al elegir una estrategia de evaluación debemos tener presente que cada técnica y/o instrumento que elijamos requerirá distintas cantidades de tiempo, personal preparado, material, puntuación y análisis (Michelson y otros, 1987). En consecuencia, será importante tener presente los recursos con los que se cuenta a fin de poder concluir una estrategia de evaluación de forma exitosa. Por lo tanto, antes de decidir una estrategia de evaluación, será necesario llevar a cabo un análisis de las necesidades que determinadas estrategias provocan para poder determinar la utilidad y posibilidad de las mismas.

d) Información que se ofrece

Al centrar nuestra atención en la información que las distintas técnicas e instrumentos de evaluación nos ofrece debemos tener presente que aquélla dependerá de

la elección que hayamos realizado. Así, a través de algunas técnicas podremos recabar información sobre el nivel de aceptación entre los compañeros (por ejemplo el sociograma), mientras que otras nos ofrecerán el grado de ejecución de determinadas habilidades (por ejemplo, la observación). Por lo tanto, al elegir una determinada estrategia de evaluación deberemos tener presente qué tipo y cantidad de información deseamos.

e) Cualidades técnicas

No basta con que utilicemos determinadas técnicas de evaluación que sean socialmente válidas, que puedan llevarse a cabo y que nos ofrezcan la información que deseamos; además, los instrumentos que utilicemos han de ser de calidad. Es decir, habrán de ser instrumentos válidos y fiables.

f) Marcos teóricos

El término *competencia social* ha sido conceptualizado de múltiples formas, reflejando cada una de ellas distintas perspectivas teóricas. Dodge, Pettit, McClaskey y Brown (1986), al analizar estas distintas concepciones llegan a la conclusión de que lo importante no es responder al interrogante de ¿qué es la competencia social?, sino intentar dar respuesta a cómo se relacionan los distintos aspectos del funcionamiento social contemplados por los diversos autores a través de sus diversas definiciones.

Responder a esta cuestión supone: (a) una formulación teórica de cómo se relacionan tales aspectos (fundamentalmente, cognitivo, conductual, reacciones de los iguales y riesgo psicológico); (b) desarrollar procedimientos y medidas para evaluar estos procesos; y, (c) estudio empírico de las relaciones actuales.

Dado que el presente trabajo se centra en la evaluación, obviaremos los otros dos aspectos, atendiendo a las implicaciones que para la evaluación tiene el modelo de intercambio social presentado por parte de Dodge y otros (Dodge, 1985; Dodge, Pettit, McClaskey y Brown, 1986). Para estos autores, el proceso de evaluación ha de tener tres fases fundamentales: (a) identificar los sujetos socialmente incompetentes, (b) determinar las situaciones sociales en las que la incompetencia social de los sujetos es un problema, y (c) determinar la fuente de valoración de la competencia social dentro de una situación determinada.

Para cada una de estas fases se proponen medios e instrumentos específicos. Así, al identificar los sujetos socialmente incompetentes, las valoraciones de otros o las técnicas sociométricas pueden ser utilizadas como medios de diagnóstico, especialmente para intervenciones preventivas; las entrevistas y la observación directa serán los instrumentos esenciales en la segunda fase; y la observación directa constituirá el medio de evaluación en la tercera y última fase.

Además de esta perspectiva general, desde el campo de la evaluación y entrenamiento de la competencia social podemos encontrarnos en la literatura actual diferentes paradigmas (Asher, 1985), que varían en los procedimientos de intervención empleados, los contenidos que se enseñan en los mismos, así como en los criterios de selección de los sujetos de entrenamiento y en las medidas de los resultados obtenidos.

Un primer paradigma, selecciona a los sujetos sobre la base de las proporciones de interacción que mantienen con sus iguales, utilizando distintos procedimientos del condicionamiento operante para la intervención (Wanlass y Prinz, 1982). Como principal característica de este paradigma cabe destacar que su atención se centra en la cantidad de interacción, olvidando otros aspectos más cualitativos de la interacción social de los sujetos.

Como contrapunto al anterior, un segundo paradigma se centra más en los aspectos cualitativos de la interacción, concretamente en la medida en que los sujetos son asertivos en sus aproximaciones y respuestas a los demás. En este caso los sujetos se seleccionan a partir de medidas en asertividad a través de autoinformes y role-play, utilizándose diversos procedimientos instructivos para incrementar la asertividad en numerosas situaciones problemáticas (Michelson, Sugai, Wood y Kazdin, 1987).

Por último, un tercer paradigma parte de la selección de los sujetos a partir de criterios sociométricos, enseñando aquellas habilidades que la investigación ha descubierto que correlaciona con el estatus sociométrico, empleando como estrategias de enseñanza la instrucción directa y procedimientos de base cognitiva en los que los sujetos son instruidos de forma verbal acerca de la importancia de distintos conceptos de la interacción social (ver Ladd y Mize, 1981, 1983).

Respecto a este último paradigma, debemos tener presente que es relativamente reciente, a pesar de lo cual se viene mostrando muy prometedor, fundamentalmente por la racionalidad con la que se realiza la selección de los sujetos, así como los resultados que se vienen obteniendo (Asher, 1985; Frenzt, Gresham y Elliot, 1991).

En definitiva, y a modo de resumen, podemos decir que la selección de una estrategia de evaluación deberemos realizarla, fundamentalmente, a partir de los interrogantes a los cuales queremos dar respuesta, y teniendo en cuenta la validez social, la practicabilidad, la información que nos aporta y la calidad de los instrumentos que utilicemos, tomando como base el marco teórico desde el que conceptualicemos la competencia social.

2. TÉCNICAS Y PROCEDIMIENTOS DE EVALUACIÓN

Como hemos señalado anteriormente, podemos decir que no existe una metodología única para llevar a cabo la evaluación de la competencia social², pudiendo encontrar distintas fuentes de variación en las diversas técnicas y procedimientos de evaluación. En este sentido Gresham (1988) presenta tres posibles fuentes de variación en los métodos de evaluación:

- a) Las fuentes de información (padres, profesores, compañeros, observadores o el propio sujeto).
- b) El grado de especificidad: Desde descripciones globales a la cuantificación de conductas específicas.
- c) La proximidad temporal en la realización de la evaluación.

Tomando en consideración estas tres fuentes de variabilidad, Elliot, Sheridan y Gresham (1989) presentan los cinco métodos fundamentales para la evaluación de las habilidades sociales que podemos observar en la Tabla I.

TABLA I
RESUMEN DE LOS MÉTODOS DE EVALUACIÓN DE LAS HABILIDADES SOCIALES Y SUS CARACTERÍSTICAS
(Elliot, Sheridan y Gresham, 1989:201)

Método	Fuente de información	Especificidad de la información	Proximidad de la conducta objeto
Sociométrico	Compañeros	Molar (estatus social)	Alejada en el tiempo
Observaciones directas	Terceras personas (psicólogo /profesor)	Molecular (conductas sociales específicas)	Alejada en el tiempo
Escalas	Profesor Padre	Molar-molecular (conducta de un dominio específico)	Alejada en el tiempo
Role-play	Uno mismo Terceras personas (psicólogo /profesor)	Molecular	Alejada en el tiempo y la situación
Entrevista conductual	Niño Profesor Padre	Molecular	Alejada en el tiempo

Por su parte, Blanco Abarca (1983) considera como los instrumentos más representativos para la evaluación de las habilidades sociales los autoinformes, la observación y las medidas fisiológicas, tal como se recoge en la Tabla II.

Alonso Tapia (1983), centrándose en la evaluación del desarrollo social, considera la existencia de cuatro modelos teóricos de evaluación, cada uno de los cuales parte de supuestos distintos y utiliza procedimientos diferentes para la evaluación (ver Tabla III).

TABLA II
TÉCNICAS DE EVALUACIÓN DE LAS HABILIDADES SOCIALES
(BLANCO ABARCA, 1983)

MODALIDADES	EJEMPLOS
Autoinformes	<ul style="list-style-type: none"> * Inventario de asertividad de Rathus (RAS) * Escala de asertividad para estudiantes (CSES) * Inventario de resolución de conflictos (CRI) * Cuestionario situacional (SQ) * Cuestionario de evitación y angustia social (SAD) * Escala de temor a la crítica negativa (FNE)
Observación en situaciones artificiales	<ul style="list-style-type: none"> * Test situacional (ST) * Test conductual de «role-playing» * Test conductual de asertividad (BAT) * Test de asertividad para estudiantes universitarias (CWAS)
Observación en situaciones naturales	<ul style="list-style-type: none"> * Test de interacción social
Medidas fisiológicas	

Por último, Gresham (1988), basándose en la medida en que la evaluación ofrece información para un análisis funcional de la conducta, presenta la clasificación de los procedimientos de evaluación atendiendo al grado en que cada una de ellas se adecúa, bien al diagnóstico o la clasificación de los sujetos, bien al proceso de intervención o terapia (ver Tabla IV).

Podemos ver, pues, que existen multitud de instrumentos para evaluar la competencia social, lo cual nos exige un análisis de los mismos a fin de orientar nuestra elección de cara al establecimiento de una posible estrategia de evaluación. A tal fin, en los siguientes apartados procederemos a una descripción de los mismos, presentando ejemplos de cada uno de ellos, así como las ventajas y limitaciones de cada uno.

2.1. Técnicas sociométricas

Las técnicas sociométricas, que requieren que cada miembro de un grupo informe de sus percepciones o sentimientos con respecto a los demás miembros del grupo, han constituido uno de los instrumentos primordiales para la identificación y clasificación de los niños con dificultades, así como una medida de la competencia

TABLA III
MODELOS TEÓRICOS Y SUPUESTOS QUE ORIENTAN LA EVALUACIÓN DEL DESARROLLO SOCIAL

MODELOS TEÓRICOS	SUPUESTOS	EJEMPLOS
PRIMER MODELO	Para cada grupo de sujetos de edad cronológica distinta hay un nivel de desarrollo social y de competencia personal que son normales. Conceptos: «Edad social» «Cociente de desarrollo social».	Escala de Madurez Social de Vineland (Doll, 1985). Medida del desarrollo psicossocial (Hurtig y Zazzo, 1969).
SEGUNDO MODELO	Igual concepción el desarrollo social que los anteriores, pero difiere en la construcción del instrumento al comparar no con sujetos normales sino sujetos de igual CI e introducir la evaluación cualitativa.	Cuadro para la evaluación del progreso en el desarrollo social (Gunzburg, 1964)
TERCER MODELO	Los diferentes comportamientos son fruto del aprendizaje. Cualquier conducta compleja puede aprenderse si se descompone en comportamientos más simples.	WUAATS (Cone, 1981).
CUARTO MODELO	Lo importante es evaluar los factores perturbadores y facilitadores de la socialización en los sujetos normales y su incidencia en otros aspectos.	Batería de socialización (Silva y Martorell, 1983).

social, además de ser tomada como medida criterio para valorar los efectos de los programas de entrenamiento. Es más, la utilización de criterios sociométricos para la selección y medida de los resultados, llega a ser considerado por Asher (1985) como un medio para identificar un paradigma en el entrenamiento de las habilidades sociales, como se reseñó anteriormente.

Previamente al análisis de las diferentes técnicas debemos tener en cuenta que aún cuando éstas son categorizadas como una metodología uniforme, comprenden diferentes procedimientos y, por otra parte, el estatus sociométrico (entendido éste como la puntuación recibida en un test sociométrico) es de carácter multideterminado, interviniendo en el mismo factores sociales y no sociales, pudiendo éstos últimos afectar al estatus sociométrico (Hops y Lewin, 1984; McConnel y Odom, 1986).

TABLA IV
MÉTODOS DE DIAGNÓSTICO DE LAS HABILIDADES SOCIALES
(GRESHAM, 1988: 533)

DIAGNÓSTICO/CLASIFICACIÓN	INTERVENCIÓN/TERAPIA
1. Técnicas sociométricas a. Nominaciones b. Escalas de valoración c. Clasificación	1. Entrevistas conductuales a. Profesor b. Padres c. Niño d. Compañeros
2. Valoraciones por otros a. Padres b. Profesores c. Compañeros	2. Observaciones directas a. Clase b. Patio de recreo c. Comedor
3. Autoinformes	3. Auto-observación con análisis funcional
4. Role Play conductual a. Análogo b. Naturalista	

Teniendo estas dos dificultades como base, las principales técnicas sociométricas que se vienen utilizando son: la nominación, la escala de valoración, la evaluación de los compañeros y la comparación por parejas.

a) Técnica de nominación

En la *técnica de nominación*, se pide a los sujetos que nombren a los compañeros que se ajustan a un determinado criterio no conductual, pudiendo realizarse nominaciones positivas o negativas, según se valore el ajuste o desajuste con relación al criterio.

Es importante tener en cuenta que el criterio utilizado en la técnica de nominación no es conductual, ya que se basa en actividades (el juego, el trabajo... etc.) o en atributos (amistad, características físicas... etc.) antes que en conductas específicas.

Dentro de este tipo de técnica podemos encontrarnos con dos instrumentos que son utilizados de forma diferente: El test sociométrico y el test de percepción sociométrica, aun cuando éste último puede considerarse como una parte del anterior.

En el test sociométrico las nominaciones se obtienen preguntando con quién se desearía jugar o trabajar (o no se desearía, en el caso de las nominaciones negativas), o solicitando que sean nombrados los mejores amigos (o aquellos niños que

menos gustan, en el caso de las nominaciones negativas). Por su parte, en el test de percepción sociométrica se pide a cada sujeto que intente acertar quién o quiénes le han elegido o rechazado, en cada una de las preguntas.

b) Escalas de Valoración

En las *escalas de valoración*, el niño valora a cada uno de sus compañeros de clase de acuerdo a un criterio determinado. Por lo general, se presenta a los niños una relación de todos los compañeros, a la derecha de la cual se ofrece una escala gradual de estimación respecto a alguna dimensión social o interpersonal, y se les pide que evalúen a cada uno de sus compañeros, pero no a sí mismos.

La principal ventaja que presenta este tipo de instrumento radica en que todos los sujetos son evaluados, a diferencia del test sociométrico, en el que las elecciones o rechazos pueden presentarse en tan sólo algunos. Pero, también tiene sus desventajas, entre las que caben citar la posible tendencia puntuar en el punto medio o evaluar de forma estereotipada (McConnel y Odom, 1986).

c) Evaluación de los compañeros

La *evaluación de los compañeros* difiere del diagnóstico sociométrico en que el primero requiere que los niños realicen juicios sobre la conducta de sus compañeros, mientras que el segundo requiere que los niños hagan juicios sobre sus sentimientos hacia los mismos. Es decir, en la evaluación de los compañeros se realizan juicios relativamente objetivos sobre un criterio determinado antes que selecciones preferenciales, como en el caso del diagnóstico sociométrico.

A través de la evaluación de los compañeros, los niños son interrogados para nombrar o clasificar a otro de acuerdo con una variedad de criterios conductuales y a través de distintas formas. Entre las distintas formas utilizadas, la más frecuente es la de las adivinanzas, a través de la cual se le ofrece al niño un listado de descripciones (por ejemplo, valiente, ordenado, el que más sabe) y el alumno «adivina» qué alumno de la clase se corresponde con el criterio. El número de elecciones es ilimitado, pudiendo cada alumno elegir a más de uno si así lo desea.

Como ejemplo de este tipo de instrumentos podemos citar el test *The Guess Who?* (Veldman y Sheffield, 1979), utilizándose también esta técnica por parte de Díaz-Aguado (1986) en su investigación sobre la interacción entre iguales. Otra modalidad de este tipo lo constituye el instrumento *Shapiro Sociometric Role Assignment Test (SSRAT)* (Shapiro y Sobel, 1981) en el que, al igual que en el *Guess Who?*, los niños leen una serie de descripciones de los compañeros y eligen al que se corresponda con cada descripción.

Un tercer procedimiento es el *Class Play* (Bower, 1960), en el que los alumnos se encargan de elegir a los compañeros que desempeñarán determinados papeles en un juego de clase, distribuyéndose los papeles en positivos (profesor, delegado de

curso) y negativos (miedoso, engreído). Una posterior revisión de este instrumento fue llevada a cabo por parte de Masten, Morison y Pellegrini (1985).

d) *Comparación por parejas*

La *comparación por parejas* es el procedimiento más laborioso y el que requiere una mayor cantidad de tiempo, lo cual explica su escaso uso. En esta técnica se toman la fotografía de todas las posibles parejas que se pueden formar con los alumnos de la clase, y el evaluador elige en cada una de ellas el miembro de la pareja que se adecúa al criterio establecido. El número de elecciones que cada alumno recibe es utilizado como una medida del estatus social.

e) *Interpretación de las puntuaciones sociométricas*

Como hemos podido comprobar al presentar cada una de estas técnicas, cada una de ellas evalúan aspectos diferentes de las relaciones que mantiene un niño con sus compañeros. En este sentido, las técnicas de nominación parecen ser una medida de la popularidad entre los compañeros, mientras que las escalas de valoración evalúan más bien el nivel global de aceptación por parte de los mismos.

Tradicionalmente, la interpretación que se ha realizado del estatus sociométrico se ha basado en las medidas globales, lo cual imposibilita una información específica. Así, en las técnicas de nominación se valoran las elecciones y los rechazos, en las escalas de valoración se suman todas las puntuaciones y en las comparaciones por parejas se suman todas las elecciones.

Actualmente se tiende hacia un análisis más específico que posibilite un mayor conocimiento de las relaciones que mantienen los niños entre sus iguales, optándose por el examen de las puntuaciones individuales, las nominaciones o valoraciones positivas mutuas de los compañeros y la estructura de las elecciones mutuas dentro de un grupo (Hallinan, 1981). En este sentido, una de las clasificaciones más interesantes es la presentada por Coie, Dodge y Coppotelli (1982), quienes llegan a identificar cinco grupos de estatus sociométrico: (1) popular, (b) ignorado, (3) rechazado, (4) controvertido, y (5) promedio. Estos grupos sociométricos se basan en las puntuaciones típicas obtenidas por los alumnos en un test sociométrico que utiliza nominaciones positivas y negativas.

Así, para cada estudiante se puede disponer de las elecciones (E) y los rechazos (R), a partir de las cuales se pueden calcular la «*preferencia social (PS)*» y el «*impacto social (IS)*». La preferencia social se calcula sustrayendo el número de rechazos al número de elecciones ($PS=E-R$); y el impacto social se calcula sumando el número de elecciones y rechazos ($IS=E+R$). Estas puntuaciones se estandarizan (Media=0, Desviación Típica=1) dentro de cada clase y los grupos de estatus sociométrico se categorizan de acuerdo a los siguientes criterios:

- a) *Grupo popular*: Incluye los estudiantes que obtienen una puntuación en preferencia social mayor de 1, una puntuación estandarizada en elecciones mayor

- que 0, y una puntuación estandarizada en rechazos menor que 0 ($PS > 1$; $E > 0$ y $R < 0$).
- b) *Grupo controvertido*: Lo constituyen los alumnos que han obtenido una puntuación en impacto social mayor que 1, y las puntuaciones estandarizadas en elecciones y rechazos son mayor de 0 ($IS > 1$, $E > 0$ y $R > 0$).
- c) *Grupo ignorado*: Constituido por los alumnos que obtienen puntuaciones en impacto social menores que -1 y las puntuaciones absolutas en elecciones de 0 ($IS < -1$, $E = 0$).
- d) *Grupo rechazado*: Incluye a los alumnos que obtienen una puntuación en preferencia social menor que -1, una puntuación estandarizada en rechazo mayor de 0 y una puntuación estandarizada en elecciones menor de 0 ($PS < -1$, $R > 0$ y $E < 0$).
- e) *Grupo promedio*: Se encuentra en el punto medio de las dimensiones preferencia social e impacto social, constituido por los alumnos que reciben una puntuación en PS mayor que -0.5 y menor que 0.5.

A partir de las escalas de valoración, Sánchez (1981) propone el cálculo de dos índices: *integración* y *cohesión*. El *índice de integración* es un valor numérico que identifica a cada sujeto del grupo dentro de una escala continua que oscila desde máxima identificación e integración en el grupo, al extremo opuesto. Para el cálculo del índice de integración de un sujeto concreto, se realiza la suma ponderada de los pronunciamientos de sus compañeros.

El *índice de cohesión*, por su parte, nos informa del grado o nivel de pertenencia que tienen los sujetos de un grupo determinado. Se calcula sumando todas las puntuaciones que los sujetos han alcanzado en la escala de valoración y dividiendo la suma total por el número de sujetos menos uno.

A pesar de que las medidas más utilizadas para evaluar la competencia social han sido las técnicas sociométricas, valorar la competencia social a partir del estatus sociométrico de un sujeto parece, cuando menos, simple y arriesgado; aun cuando supone un primer elemento importante del proceso de evaluación.

«Las medidas sociométricas tienen características deseables que apoyan fuertemente su utilización en el diagnóstico de la competencia social de los niños. No obstante, otras características de estas medidas sugieren la necesidad de ser cautelosos al interpretar los resultados de los diagnósticos sociométricos. De forma global, no podemos concluir que el estatus sociométrico es el único criterio, o necesariamente el mejor, para evaluar la competencia social de los niños. Más bien concluimos que las medidas sociométricas son componentes necesarios, pero no suficientes, de un modelo de medida múltiple para el diagnóstico de la competencia social». (McConnel y Odom, 1986:269).

2.2. Valoraciones por otros

Entre los métodos más utilizados para la evaluación de la competencia social podemos considerar las valoraciones realizadas por parte de compañeros, padres y profesores, siendo éstas últimas las que se han utilizado con más frecuencia (McMahon, 1984).

Las valoraciones del profesorado se viene obteniendo a través de tres procedimientos fundamentales: (a) La *nominación*, a través de la cual se pide al profesor que especifique los niños que se ajustan a un determinado criterio (por ejemplo, agresivo, tímido, responsable); (b) los *cuestionarios*, en los que el profesor indica los descriptores que reflejan la conducta de cada niño y, (c) la *escalas de apreciación*, en las que el profesor estima de forma cuantitativa el grado en que un descriptor determinado es característico de cada niño.

Al igual que otras técnicas, las valoraciones de los profesores presentan sus ventajas e inconvenientes (French y Tine, 1982). Entre las primeras podemos destacar la facilidad con que se recogen los datos, así como la ventaja adicional de que el profesor está en contacto continuo con todos y cada uno de los alumnos, teniendo acceso a comparaciones normativas cuando toma decisiones sobre la naturaleza atípica de la conducta del niño.

Entre los inconvenientes podemos destacar que, por una parte, aun cuando existe correlación significativa entre los datos aportados por las valoraciones de los profesores y las técnicas sociométricas, la magnitud de esta coincidencia no es del todo convincente. Por otra parte, la valoración de los profesores, a veces, se basa en criterios de selección inadecuados, partiendo de supuestos estereotipados sobre la conducta inadaptativa, incorporando sus criterios a las línea de referencia.

Como ejemplos de este tipo de instrumentos, podemos señalar los siguientes:

a) *Matson Evaluation of Social Skills with Youngers (MESSY)*.

Esta escala, desarrollada por Matson, Rotatori y Helsel (1983), consta de un total de 64 ítems, cada uno de los cuales es valorado por parte del profesor en una escala de cinco puntos (1=nunca, 5=mucho). Los 63 ítems se agrupan en torno a dos factores: asertividad inapropiada/impulsividad y habilidades sociales apropiadas. Respecto a las propiedades psicométricas, en las investigaciones preliminares presenta una fiabilidad test-retest de 0,95, pudiendo discriminar entre la severidad de los problemas manifestados por los niños.

b) *Social Behavior Assessment (SBA)*

Desarrollada esta escala por parte de Stephens (1979), se trata de una escala de carácter comprensivo con un total de 136 ítems presentados en una escala de tipo Lickert agrupados en torno a cuatro factores: ambiente, interpersonal, relacionado consigo mismo y relacionado con tareas.

c) *Taxonomy of Problematic Social Situations (TOPS)*

Esta escala, desarrollada por Dodge, McClaskey y Feldman (1985), consta de 44 ítems que son evaluados por parte del profesor sobre una escala de cinco puntos (1=nunca, 5=casi siempre). Los seis factores que se contemplan en la escala son: entrada en el grupo de iguales, respuesta a las provocaciones de los compañeros, respuesta al fracaso, respuesta al éxito, expectativas sociales y expectativas del profesor.

d) *Teacher Rating of Social Skills (TROSS)*

Esta escala, contruida por Clark, Gresham y Elliot (1985), consta de 50 ítems en la que el profesor clasifica la frecuencia de la conducta social sobre la base de una escala de tres puntos (con frecuencia es verdad, algunas veces es verdad, nunca es verdad). En cada una de las conductas el profesor ha de reseñar la importancia que cada una de ellas tiene para el éxito en la clase, incrementándose de esta forma la validez social.

e) *Walker Social Skills Curriculum Scale*

La escala diseñada por Walker et al (1983), consta de un total de 28 ítems ordenados por el profesor en una escala de cinco puntos. Con posterioridad, Walker y McConnel (1988), han desarrollado una nueva escala, *Walker-McConnel Scale of Social Competence and School Adjustment*, que consta de un total de 43 ítems formulados positivamente en una escala Lickert de 5 puntos (desde 1=nunca ocurre, hasta 5=frecuentemente ocurre), siendo diseñada para medir la percepción del profesor de la competencia social y el ajuste escolar del alumno. A través del análisis factorial, en esta escala se han identificado tres subescalas: (a) Conducta social preferida por el profesor, (b) Conducta social preferida por los compañeros y (c) Conducta de ajuste escolar.

f) *Escalas ESE-1 y ESE-2. (Pelechano, 1979)*

La *Escala ESE-1 de hábitos positivos de socialización en ambientes escolares*, consta de un total de 100 ítems, cada uno de los cuales tiene cuatro alternativas de respuesta (A= nunca, D= siempre), distribuidos estos ítems en cuatro factores: (a) Colaboración con el profesor y los compañeros, (b) respeto hacia los demás y sí mismo y responsabilidad en lo que dice y hace, (c) popularidad y liderazgo, y (d) actitud diferencial positiva entre tareas escolares y extraescolares aparejado a responsabilidad ante sí mismo y los mayores.

Por su parte, la *Escala ESE-2 de hábitos perturbadores de la socialización en ambientes escolares*, consta de un total de 87 ítems, con cuatro posibles modalidades de respuesta, al igual que la ESE-1, y distribuidos en dos factores: (a) Hiperactividad y agresividad, y (b) inseguridad y retraimiento.

g) *Batería de Socialización (Silva y Martorell, 1983)*

Tomando como base las escalas ESE-1 y ESE-2, Silva y Martorell (1983) desarrollan la Batería de Socialización BAS-1, BAS-2, para profesores y padres, respectivamente.

Estas escalas de estimación, para ser aplicadas a niños y adolescentes de 6 a 15 años, constan de cuatro dimensiones facilitadoras de la socialización (Liderazgo, Jovialidad, Sensibilidad social y Respeto-Autocontrol), tres perturbadoras de la misma (Agresividad-terquedad, Apatía-retraimiento y Ansiedad-timidez), y una escala global de adaptación social o Criterial-socialización.

h) *Social Skills Rating System (SSRS) (Gresham y Elliot, 1990)*

El SSRS consta de un manual, cuestionarios para alumnos, profesores y padres, un registro para el diagnóstico y la intervención, así como un programa informático para la corrección e interpretación de la SSRS.

Los cuestionarios para padres y profesores son los que podemos considerar como valoraciones de otros, en cuanto a través de los mismos, en una escala de 0 a 2, se expresa la frecuencia con que el alumno realiza determinadas conductas, así como la importancia que esa conducta tiene. Las escalas y subescalas varían en función del nivel (preescolar, primaria y secundaria), aun cuando podemos señalar que se centra en la valoración de tres dimensiones: Habilidades sociales, problemas conductuales y competencia académica.

2.3. Métodos de clasificación

Los métodos de clasificación, especialmente por parte de los profesores, han sido utilizados frecuentemente para la selección de los niños que serían objeto de entrenamiento en habilidades sociales (Gresham, 1986). Desde esta perspectiva, los profesores clasifican a los alumnos por orden de acuerdo con algún criterio, pudiendo éste ser de carácter conductual (el que menos habla, el más agresivo... etc.) o no conductual (el más querido, el que tiene menos amigos... etc.).

Como principal limitación de este tipo de instrumentos nos encontramos con que los profesores pueden fallar al identificar a los niños socialmente introvertidos, pues las conductas de estos alumnos no resultan problemáticas para los profesores (Hops y Greenwood, 1981). Por otra parte, los resultados obtenidos con este tipo de instrumentos mantienen niveles bajos de concordancia entre distintos evaluadores. Así, podemos encontrarnos con que las valoraciones realizadas entre padres, profesores y compañeros, sobre un mismo niño, difieran enormemente.

2.4. Juego de Papeles Conductual (Behavioral Role Play)

Esta técnica se viene utilizando, fundamentalmente, por las ventajas que presenta ante las técnicas sociométricas, las escalas de valoración y la observación en situaciones naturales, entre las que caben destacar (Gresham, 1986):

a) Puede diagnosticar importantes conductas sociales que ocurren con poca frecuencia en situaciones naturales.

b) Supone la representación de una habilidad, más que la valoración o percepción de esa habilidad.

c) Las situaciones simuladas pueden estar más controladas para diagnosticar la respuesta de un niño al seleccionarse los estímulos.

d) Se pueden reconstruir ambientes en los que resulta difícil observar la conducta del niño.

e) Es menos costoso que la recogida de datos a través de la observación en lugares naturales.

A pesar de estas ventajas, el juego de Papeles presenta una seria limitación, como es la poca correlación con la observación en situaciones naturales, así como su escaso nivel predictivo del estatus sociométrico, lo cual lleva a considerar a esta técnica válida para el diagnóstico en la ejecución de habilidades sociales, antes que una técnica de diagnóstico para la intervención o la evaluación de los resultados de los programas de intervención.

2.5. Autoinformes

A través de los autoinformes (self-report), los propios sujetos muestran sus opiniones, actitudes y sentimientos (Finch y Rogers, 1984), constituyéndose en instrumentos que también han venido utilizándose para la evaluación de las habilidades sociales. No obstante, presentan una serie de deficiencias que les hace ser considerados, por parte de algunos autores, como poco útiles (Maag, 1989).

«...las medidas de autoinforme de los niños no parecen ser útiles al predecir la aceptación de los compañeros, la popularidad entre los compañeros, las clasificaciones del profesorado de las habilidades sociales, la ejecución en el role-play, o la conducta social en situaciones naturales.» (Gresham, 1986:103).

A pesar de estas deficiencias, a partir del autoinforme se podrán tomar datos de las percepciones que los propios sujetos tienen sobre sí mismos, y los cambios que los programas de entrenamiento hayan podido producir en tales percepciones.

Como ejemplo de autoinforme podemos citar el elaborado por Asher y Wheeler (Asher, 1985), para analizar si los niños con diferentes estatus sociométrico diferían en sus sentimientos de soledad e insatisfacción social. Del total de 24 ítems, dieci-

séis de ellos se centran en los sentimientos de soledad de los niños (me siento sólo en el colegio), sentimientos de adecuación/inadecuación social (soy bueno trabajando con otros compañeros de la clase). Los ocho ítems restantes fueron incluidos por parte de los autores para ayudar al niño a sentirse más abierto y relajado. Los niños responden a cada uno de los ítems sobre una escala de cinco puntos sobre el grado en que cada uno de ellos dice la verdad sobre sí mismo.

También podemos reseñar como ejemplo de autoinforme las versiones para alumno de la *Social Skills Rating System* de Gresham y Elliot (1990), para enseñanza primaria y secundaria.

Otros ejemplos de autoinforme lo constituyen el *Inventario de Adaptación Social* de Pérez Juste (1983), el *Cuestionario de Adaptación Escolar* de Jiménez Fernández (1979) y la *Batería de Socialización BAS-3*, construida esta última por parte de Silva y Martorell (1989) para su aplicación a sujetos de 11 a 19 años, a partir de las escalas de estimación BAS-1 y BAS-2, y con el objetivo de indagar sobre la percepción que los propios sujetos tienen de su conducta social, evaluada ésta a través de las siguientes dimensiones: Consideración con los demás, autocontrol en las relaciones sociales, retraimiento social, ansiedad social/timidez y liderazgo.

2.6. Entrevista conductual

Cuando hablamos de entrevista, debemos tener presente que ésta exige unos requisitos, a saber (Silva, 1983):

- Una relación directa entre personas (dos o más de dos).
- Una vía de comunicación simbólica, preferentemente oral.
- Unos objetivos prefijados y conocidos, al menos, por el entrevistador.
- Una asignación de roles que significa un control de la situación por parte del entrevistador.

A pesar del poco uso que se hace de la entrevista conductual de cara al diagnóstico de la competencia social de los niños, constituye un instrumento muy útil para definir las conductas sociales en términos observables, identificando las condiciones antecedentes, secuenciales y consecuentes que afectan al desarrollo, y para diseñar sistemas de observación para medir objetivos conductuales (Gross, 1984; Elliot y otros, 1987); y, al igual que otras técnicas, presenta sus ventajas e inconvenientes, entre los que cabe destacar (Silva, 1983):

a) Ventajas:

- Instrumento flexible, tanto en su desarrollo como en la temática a la que se aplique.
- Provee simultáneamente indicadores tanto verbales como no verbales de la respuesta, y de su mutua relación.
- Establece las bases y la oportunidad para una relación interpersonal dialogal.

- Ofrece la posibilidad de un intercambio confidencial de información.
 - Posibilita la retroalimentación, modificando las respuestas de entrevistador y entrevistado según las necesidades.
 - Sirve como alternativa a las personas impedidas de responder por escrito, o a las situaciones no accesibles a la observación directa.
- b) Inconvenientes:
- Es un instrumento indirecto de recopilación de información.
 - Existe poca investigación en torno a la fiabilidad y validez.
 - No existen procedimientos tipificados para la codificación, interpretación y estrategias de decisión.
 - Exige una gran cantidad de tiempo para su realización.
 - Es preciso contar con personal preparado.

2.7. Observación directa

La observación directa de las interacciones sociales que mantiene un sujeto en su ambiente natural es considerado por algunos autores como el medio ideal para llevar a cabo la evaluación de las habilidades sociales (Michelson, 1987; Elliot y otros, 1987), dado su carácter directo y ecológico, pudiéndose utilizar tanto en la identificación de los sujetos objeto de tratamiento como para la evaluación de los resultados de la intervención.

«La observación directa de la conducta está en el corazón del diagnóstico de la conducta del niño». (Barton y Ascione, 1984:167).

De forma genérica, la observación se lleva a cabo con uno o más observadores que registran la frecuencia, duración y/o calidad de los comportamientos que se han identificado como socialmente válidos e importantes, y tras un proceso de operativización de tales comportamientos, a fin de facilitar una observación fiable y válida, resultando su validez externa su gran virtud.

A pesar de la riqueza que presenta esta técnica, existen ciertas limitaciones metodológicas (Michelson, 1987), como pueden ser el sesgo de la expectativa entre observadores, la complejidad del sistema y el conocimiento de la fiabilidad de la evaluación, además de otros problemas tales como la practicabilidad, dada la necesidad de contar con personal formado, así como las distintos niveles de información que pueden darse.

2.8. Auto-observación

La auto-observación consta de un doble proceso consistente, por un lado, en atender a la propia conducta y, por otro, en registrarla mediante algún procedimien-

to establecido con anterioridad (Avia, 1983), pudiendo establecerse desde un diario en el que el sujeto registre toda la información sin límite alguno, hasta la clasificación de la ocurrencia de determinadas conductas, pensamientos y sentimientos, previamente operativizados (Shapiro, 1984).

Como principal inconveniente de esta técnica cabe reseñar la naturaleza subjetiva de la información obtenida (Maag, 1989), aun cuando a través de esta subjetividad puede detectarse, al contrastar con otras técnicas, las dificultades que el sujeto presenta en sus habilidades socio-cognitivas.

3. PERSPECTIVAS DE INVESTIGACIÓN

Hasta aquí hemos presentado una revisión de las diferentes técnicas e instrumentos que se vienen utilizando en la evaluación de la competencia social. Asumir esta evaluación nos abre toda una serie de posibles vías de investigación, que esperamos y deseamos que poco a poco se vayan cubriendo, entre las cuales podemos destacar:

a) *Delimitación del constructo «competencia social»*. Siempre que procedemos a la evaluación debemos tener muy claro qué queremos evaluar, y como hemos señalado en otro lugar (García Jiménez, García Pastor y Rodríguez Gómez, en prensa), el propio constructo «competencia social» aparece a lo largo de la literatura insuficientemente delimitado. Las propias características, peculiaridades y dimensiones o variables que contemplan los múltiples medios e instrumentos de evaluación que se están utilizando actualmente nos vienen a confirmar esta pluralidad en la conceptualización de la competencia social.

b) *Propuesta y validación de modelos de evaluación de la competencia social*. A lo largo de las páginas precedentes hemos podido comprobar el escaso número de instrumentos existentes en nuestro contexto socio-cultural, y estos pocos, a su vez, exigen en su mayoría una actualización no sólo terminológica sino conceptual. Mas la construcción de instrumentos de evaluación no ha de afrontarse sólo por la escasez de los mismos, sino, y esto es lo fundamental, porque sólo en la medida en que se dispongan de buenos instrumentos de evaluación podremos avanzar en la validación de constructo de la propia competencia social.

Al planificar estos modelos de evaluación, deberíamos contemplar, en primer lugar, el propio alumno, considerando las relaciones que mantiene con otros compañeros, su capacidad de autorregulación, su nivel de ajuste escolar, y el tipo de conducta manifestada en sus actuaciones, dimensiones todas estas que tienen una gran importancia, como lo demuestran las investigaciones realizadas en este campo como las de Gresham (1988), Elliot, Sheridan y Gresham (1988), Frenztz, Gresham y Elliot (1991), Ladd y Mize (1983) y Wentzel (1991a), entre otras.

Por otra parte, deberíamos valorar el contexto inmediato (fundamentalmente centrado en el clima de aula y centro), a fin de determinar las características del

contexto que rodea al alumno, dada la importancia del mismo en cuanto que condiciona las relaciones entre alumnos y profesores, facilitando o interfiriendo en el desarrollo social de los alumnos (Fuentes Vicente, 1986; Rodríguez Gómez, 1991), aspectos sobre los cuales pueden consultarse a modo de introducción los trabajos de Alonso Tapia et al. (1991), Hoy y Clover (1986), Hoy, Tarter y Bliss (1990), López (1990), Medina (1988) y Medina y Sevillano (1991), entre otros.

Mas la construcción de modelos de evaluación, desde una perspectiva educativa, no ha de quedar al margen del proceso global de intervención. En consecuencia, no deberíamos gastar los esfuerzos en llevar a cabo la construcción de instrumentos aislados, sino insertos en un modelo de evaluación global, en el que el centro sea el objeto de la evaluación y no los medios que se utilizan para la misma.

c) *Diseñar, ejecutar y valorar programas de intervención que persigan un mayor desarrollo de la competencia social de los alumnos.* De poco o nada nos serviría la evaluación si no la entendemos en un proceso de intervención educativa cuyo objetivo último sea el desarrollo del alumno. Lo que realmente interesa es potenciar la competencia general de los alumnos, su excelencia personal y, en consecuencia, el proceso de evaluación no es más que una parte del proceso de intervención educativa, evitando de esta forma caer en una evaluación que acabe en sí misma y nos conduzca a un proceso de *etiquetado* (García Pastor, 1993) que en nada ayudaría al alumno.

Ahora bien, esta necesidad de intervención exige no sólo una adaptación de lo que se viene haciendo en otros países, sino la creación y valoración de programas a partir de las características idiosincráticas de nuestro contexto. Hasta la fecha la casi totalidad de programas diseñados a este fin provienen del contexto clínico-terapéutico, que adolece de serias deficiencias de cara a su aplicabilidad, aspecto éste que hemos analizado en otros trabajos y sobre el que hemos planteado algunas posibles alternativas (G. Jiménez, G. Pastor y Rodríguez Gómez, en prensa), sírvanos tan sólo en este momento recordar la necesidad de acometer un proceso de intervención educativa desde una perspectiva ecológica de forma tal que en el currículum se explicita de forma clara el desarrollo de las relaciones sociales en la escuela.

REFERENCIAS BIBLIOGRÁFICAS

- ALONSO TAPIA, J. (1983): Evaluación del desarrollo intelectual y social. En R. Fernández Ballesteros y J. A. I. Carrobes. *Psicodiagnóstico*. Madrid: UNED.
- ALONSO TAPIA, J. et al. (1991): Sistemas motivacionales en el aula: El cuestionario CMC-2. *Boletín del Instituto de Ciencias de la Educación UAM*, 18, 51-66.
- ASHER, S. R. (1985): An Evolving Paradigm in Social Skills Training with Children. En B. H. Schneider, K. H. Rubin y J. E. Ledingham (Eds.): *Children's Peer Relations: Issues in assessment and Intervention*. New York: Springer-Verlag.
- AVIA ARANDA, M. D. (1983): La auto-observación. En R. Fernández Ballesteros y J. A. I. Carrobes. *Evaluación conductual*. Madrid: Pirámide.

- BARTON, E. J. y ASCIONE, F. R. (1984): Direct Observation. En T. H. Ollendick y M. Hersen. *Child Behavioral Assessment*. New York: Pergamon Press.
- BLANCO ABARCA, A. (1983): Evaluación de las habilidades sociales. En R. Fernández Ballesteros y J. A. I. Carrobes. *Psicodiagnóstico*. Madrid: UNED.
- BOWER, E. M. (1960): *Early Identification of Emotionally Handicapped Children in School*. Springfield, IL: Charles C. Thomas.
- CASANOVA, M. A. (1991): *La sociometría en el aula*. Madrid: La Muralla.
- CLARK, L.; GRESHAM, F. M. y ELLIOT, S. N. (1985): Development and Validation of a Social Skills Assessment Measures: The TROSS-C. *Journal of Psychoeducational Assessment*, 4, 347-356.
- COIE, J. D., DODGE, K. A. y COPOTELLI, H. (1982): Dimensions and Types of Social Status: A Cross-age Perspective. *Developmental Psychology*, 18, 557-570.
- COIE, J. D. y KREHBIEL, G. (1984): Effects of Academic Tutoring on the Social Status of Low-achieving, Socially Rejected Children. *Child Development*, 55, 1.465-1.478.
- DÍAZ-AGUADO, M. J. (1986): *El papel de la interacción entre iguales en la adaptación escolar y el desarrollo social*. Madrid: CIDE.
- DODGE, K. A. (1983): Behavioral Antecedents of Peer Social Status. *Child Development*, 54, 1.386-1.399.
- DODGE, K. A. (1985): Facets of Social Interaction and the Assessment of Social Competence in Children. En B. H. Schneider, K. H. Rubin y J. E. Ledingham. *Children's Peer Relations: Issues in assessment and Intervention*. New York: Springer-Verlag.
- DODGE, K. A.; MCCLASKEY, C. L. y FELDMAN, E. (1985): Situational Approach to the Assessment of Social Competence in Children. *Journal of Consulting and Clinical Psychology*, 53, 344-353.
- DODGE, K. A.; PETTIT, G. S.; MCCLASKEY, C. L. y BROWN, M. M. (1986): Social Competence in Children. *Monographs of the Society for Research in Child Development*, 51, (2).
- ELLIOT, S. N.; GRESHAM, F. M. y HEFFER, R. W. (1987): Social Skills Interventions: Research Findings and Training Techniques. En C. A. Mahers y J. E. Zins (Eds.): *Psychoeducational interventions in the schools*. New York: Pergamon.
- ELLIOT, S. N.; SHERIDAN, S. N. y GRESHAM, F. M. (1989): Assessing Treating Social Skills Deficits: A Case Study for Scientist-Practitioner. *Journal of School Psychology*, 27, (2), 197-222.
- FINCH, A. J. y ROGERS, T. R. (1984): Self-Report Instruments. En T. H. Ollendick y M. Hersen: *Child Behavioral Assessment*. New York: Pergamon Press.
- FRENCH, D. C. y TYNE, T. F. (1982): The identification and Treatment of Children with Peer-Relationship Difficulties. En J. P. Curran y P. M. Monti (Eds.): *Social Skills Training*. New York: The Guilford Press.
- FRENTZ, C.; GRESHAM, F. M. y ELLIOT, S. N. (1991): Popular, Controversial, Neglected, and Rejected Adolescents: Contrasts of Social Competence and Achievement Differences. *Journal of School Psychology*, 29, 109-120.
- FUENTES VICENTE, A. (1986): *Procesos funcionales y eficacia de la escuela*. Tesis Doctoral. Madrid: Universidad Complutense.
- GARCÍA JIMÉNEZ, E.; GARCÍA PASTOR, C. y RODRÍGUEZ GÓMEZ, G. (En prensa): Limitaciones del constructo «Habilidades Sociales» para la elaboración de un modelo de intervención social en el aula. *Revista de Enseñanza*.
- GARCÍA PASTOR, C. (1993): *Una escuela común para niños diferentes: La integración escolar*. Barcelona: PPU.
- GIRALDO, A. (1992): *Relaciones psicosociales en las aulas integradas*. Comunicación presentada en el X Congreso Nacional de Pedagogía (paper).

- GRESHAM, F. M. (1986): Conceptual Issues in the Assessment of Social Competence in Children. En P. S. Strain, M. J. Guralnick y H. M. Walker. *Children's Social Behavior*. New York: Academic Press.
- GRESHAM, F. M. (1988): Social Skills. Conceptual and Applied Aspects of Assessment, Training, and Social Validation. En J. C. Witt, S. N. Elliot y F. M. Gresham, F. M. *Handbook of Behavior Therapy in Education*. New York: Plenum Press.
- GRESHAM, F. M. y ELLIOT, S. N. (1990): *Social Skills Rating System*. Circles Pines, MN: American Guidance Service.
- GROSS, A. M. (1984): Behavioral Interviewing. En T. H. Ollendick y M. Hersen: *Child Behavioral Assessment*. New York: Pergamon Press.
- HOPS, H. y LEWIN, L. (1984): Peer Sociometrics Forms. En T. H. Ollendick y M. Hersen. *Child Behavioral Assessment*. New York: Pergamon Press.
- HOPS, H. y GREENWOOD, C. R. (1981): Social Skills Deficits. En E. J. Mash y L. G. Terdal (Eds.): *Behavioral assessment of childhood disorders*. New York: The Guilford Press.
- HOY, W. K. y CLOVER, S. (1986): Elementary School Climate: A Revision of the OCDQ. *Educational Administration Quarterly*, 22(1), 93-110.
- HOY, W. K.; TARTER, C. J. y BLISS, J. R. (1990): Organizational Climate, School Health, and Effectiveness: A Comparative Analysis. *Educational Administration Quarterly*, 26(3), 260-279.
- JIMÉNEZ FERNÁNDEZ, C. (1979): *Cuestionario de Adaptación Escolar*. Madrid: CSIC-Instituto de Pedagogía «San José de Calasanz».
- LADD, G. W. (1981): Effectiveness of social learning method for enhancing children's social interaction and peer acceptance. *Child Development*, 52, 171-178.
- LADD, G. W. y MIZE, J. (1983): A Cognitive-social Learning Model of Social-Skill Training. *Psychological Review*, 90, (2), 127-157.
- LÓPEZ YÁÑEZ, J. (1990): *Formación de líderes escolares mediante el ordenador*. Sevilla: GID.
- MAAG, J. W. (1989): Assessment in Social Skills Training: Methodological and Conceptual Issues for Research and Practice. *RASE*, 10, (4), 6-17.
- MASTEN, A. S., MORISON, P. y PELLEGRINI, D. (1985): A revised class play method of peer assessment. *Developmental Psychology*, 21, 523-533.
- MATSON, J. L.; ROTATORI, A. F. y HELSEL, W. J. (1983): Development of a Rating Scale to Measure Social Skills in Children: The Matson Evaluation of Social Skills with Youngers (MES-SY): *Behavior Research and Therapy*, 21, 335-340.
- MCCONNELL, S. R. y ODOM, S. L. (1986): Sociometrics: Peer-Referenced Measures and the Assessment of Social Competence. En P. S. Strain, M. J. Guralnick y H. M. Walker: *Children's Social Behavior*. New York: Academic Press.
- MCMAHON, R. J. (1984): Behavioral Checklists and Rating Scales. En T. H. Ollendick y M. Hersen *Child Behavioral Assessment*. New York: Pergamon Press.
- MEDINA RIVILLA, A. (1988): *Didáctica e interacción en el aula*. Madrid: Cincel.
- MEDINA, A. y SEVILLANO, M. L. (1991): *El clima sociorelacional en el aula*. Madrid: UNED.
- MICHELSON, L.; SUGAI, D. P.; WOOD, R. P. y KAZDIN, A. E. (1987): *Las habilidades sociales en la infancia*. Barcelona: Martínez Roca.
- PARKER, J. G. y ASHER, S. R. (1987): Peer relations and later personal adjustment: Are low-accepted children at risk? *Psychological Bulletin*, 102, 357-389.
- PELECHANO, V. (1979): *Psicología educativa comunitaria*. Valencia: Alfaplus.
- PÉREZ JUSTE, R. (1979): *Inventario de Adaptación Social*. Madrid: CSIC-Instituto de Pedagogía «San José de Calasanz».
- PÉREZ JUSTE, R. (1991): *Evaluación de Programas*. I Simposio Nacional sobre Programas de Enseñar a Pensar. Granada, julio, (paper).

- PUTALLAZ, M. (1983): Predicting children's sociometric status from their behavior. *Child Development*, 54, 1.417-1.426.
- RODRÍGUEZ GÓMEZ, G. (1991-1992): *Proyecto de investigación para el desarrollo de un programa de habilidades sociales*. Documento inédito.
- RODRÍGUEZ GÓMEZ, G. (1991): *Investigación evaluativa en torno a los factores de eficacia escolar de los centros públicos de EGB*. Tesis Doctoral. Madrid: UNED.
- SÁNCHEZ SÁNCHEZ, S. (1981): *La tutoría en los centros docentes*. Madrid: Escuela Española.
- SHAPIRO, E. S. (1984): Self-Monitoring Procedures. En T.H. Ollendick y M. Hersen. *Child Behavioral Assessment*. New York: Pergamon Press.
- SHAPIRO, S. B. y SOBEL, M. (1981): Two Multinomial Random Sociometric Voting Models. *Journal of Educational Statistics*, 6, 287-310.
- SILVA MORENO, F. (1983): La entrevista. En R. Fernández Ballesteros y J. A. I. Carroble. *Evaluación conductual*. Madrid: Pirámide.
- SILVA, F. y MARTORELL, M. C. (1983): *BAS-1, 2. Bateria de socialización (para profesores y padres)*. Madrid: TEA.
- SILVA, F. y MARTORELL, M. C. (1989): *BAS-3 (Autoevaluación)*. Madrid: TEA.
- STEPHENS, T. M. (1979): *Social Behavior Assessment*. Columbus, OH: Cedars Press.
- TERRY, R. y COIE, J. D. (1991): A Comparison of Methods for Defining Sociometric Status Among Children. *Developmental Psychology*, 27,(5), 867-880.
- VELDMAN, D. J. y SHEFFIELD, J. R. (1979): The scaling of sociometrics nominations. *Educational and Psychological Measurement*, 39, 99-106.
- WALKER, H. M.; MCCONNELL, S. R.; WALKER, J.; HOLMES, D.; TODIS, B. y GOLDEN, N. (1983): *ACCEPTS: A curriculum for effective peer and teacher skills*. Austin, TX: PRO-ED.
- WALKER, H. M. y MCCONNELL, S. (1988): *The Walker-McConnell Scale of Social Competence and School Adjustment*. Atustin, TX: PRO-ED.
- WANLASS, R. L. y PRINZ, R. J. (1982): Methodological Issues in Conceptualizing and Treating Social Isolation. *Psychological Bulletin*, 92, 35-55.
- WENTZEL, K. R. (1991a): Social Competence at School: Relations Between Social Responsibility and Academic Achievement. *Review of educational Research*, 61, (1), 24.
- WENTZEL, K. R. (1991b): Relations between Social Competence and Achievement in Early Adolescence. *Child Development*, 62, 1.066-1.078.

TRABAJOS METODOLÓGICOS

Revista Investigación Educativa - N.º 21 - 1993 (P. 107-124)

MODELOS BORROSOS EN LA MEDICIÓN ESCOLAR

por

Xavier Gil Quesada y Joan Mateo Andrés

RESUMEN

Este artículo pretende dar una breve descripción de la TCB que permita una mejor comprensión de nuestra propuesta de modelos borrosos para la medición educacional, concretada en

- dos modelos descriptivos: el conjunto de respuestas a un test y la zona de corte borroso, y
- dos modelos analíticos ligados al estudio de las estructuras de homogeneidad y jerarquización de un test.

ABSTRACT

This paper tries to offer a concise description about the Fuzzy Sets Theory which allows a better understanding of our fuzzy models offer for educational measurement based in

- two descriptive models: the set of a test answers and the fuzzy cutoff zone, and
- two analitic models connected to the study of homogeneity and hierarchy structures of a test.

La teoría clásica de conjuntos fundamenta la construcción de modelos (extremadamente fecundos) en la práctica totalidad de las áreas científicas. Esta teoría está asociada a la lógica booleana que se caracteriza por los principios del tercero excluido y el de no contradicción.

Pero esta lógica no ha posibilitado el tratamiento de aspectos importantes de la realidad que son frecuentes en la vida ordinaria y básicos en las Ciencias Humanas.

Existen universos de discurso en los que encontramos lo que la teoría clásica denomina «conjuntos mal definidos»: hombres jóvenes, personas enfermas, alumnos con un nivel de aprendizaje suficiente, etc. Son estos conceptos, que la Teoría Clásica de Conjuntos deja sin fundamentación, los que son posibles tratar a partir de la Teoría de los Conjuntos Borrosos (TCB).

En particular la TCB se muestra muy adecuada en muchos dominios relacionados con las Ciencias de la Educación y en especial en el área de la medida escolar. Este artículo pretende dar una breve descripción de la TCB que permita una mejor comprensión de nuestra propuesta de modelos borrosos para la medición educacional, concretada en

- dos modelos descriptivos: el conjunto de respuestas a un test y la zona de corte borroso, y
- dos modelos analíticos ligados al estudio de las estructuras de homogeneidad y jerarquización de un test.

Nota: una exposición más detallada y extensa de estos modelos puede encontrarse en Gil Quesada, X. («La Teoría de los Conjuntos Borrosos y la medición escolar», Tesis Doctoral, 1990).

1. BREVE RESEÑA DE LA TEORÍA DE LOS CONJUNTOS BORROSOS

En la teoría clásica de conjuntos, un subconjunto A del referencial X está bien definido si y sólo si existe un criterio que permite decidir cuándo un elemento cualquiera de X pertenece o no al conjunto A.

Es decir, todo subconjunto A establece una división dicotómica en el conjunto de los elementos de referencia. Dado $x \in X$, se verifica que

- o bien x es un elemento de A,
- o bien x no es un elemento de A,

y no hay ninguna otra posibilidad.

Por tanto, todo subconjunto A de X viene definido por una aplicación

$$f_A: X \longrightarrow \{0,1\}$$

que asigna el valor 1 a los elementos de A y 0 al resto y que se denomina *función característica del conjunto A*.

En 1965, L. A. Zadeh («Fuzz y sets», Information and control) introdujo la Teoría de los Conjuntos Borrosos (TCB) partiendo de una generalización de la función característica. La idea central es definir los subconjuntos A del referencial X mediante una *función de pertenencia*

$$\begin{array}{l} f_A: X \longrightarrow [0,1] \\ x \longrightarrow f_A(x) \end{array}$$

que generaliza la función característica de los subconjuntos clásicos (o nítidos); de este modo, dado un elemento $x \in X$, no sólo se pueden dar los dos casos

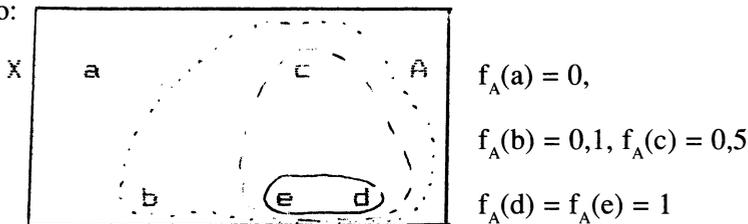
$$f_A(x) = 0 \text{ y } f_A(x) = 1$$

sino que aparecen situaciones en que

$$f_A(x) = t, t \in (0,1)$$

el elemento x tiene un grado de pertenencia t al conjunto A que no es ni cero ni uno, sino un valor intermedio.

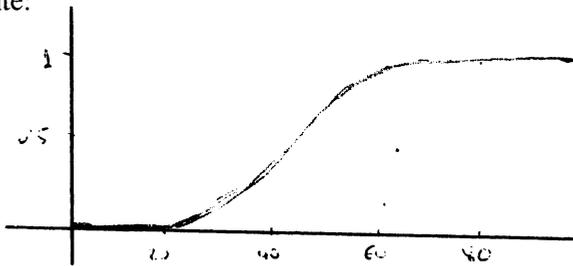
Ejemplo:



Nota: para simplificar la notación $f_A(x)$ se indica $A(x)$.

Cuando el conjunto con el que tratemos sea infinito la representación de un subconjunto borroso vendrá dado necesariamente por su función de pertenencia.

Ejemplo: el subconjunto de los hombres viejos podría definirse por una función como la siguiente:



Al caracterizar los subconjuntos borrosos A de X mediante funciones de pertenencia, son éstas las que permiten definir los conceptos básicos de la Teoría (igualdad, inclusión, reunión, intersección y paso al complementario).

En el conjunto de las partes borrosas de X se cumplen las propiedades algebraicas que definen una estructura de Álgebra de Morgan; pero, a diferencia de lo que ocurre con el complementario en la teoría clásica, no tiene por qué ocurrir que $A \cap \bar{A} = \emptyset$ ni tampoco que $A \cup \bar{A} = X$.

La teoría de Zadeh rompe la dicotomía entre el «pertenece y el no pertenece» de la teoría clásica de conjuntos al introducir una relación de pertenencia gradual:

si $A(x) \geq \alpha$, podemos decir que x pertenece al nivel α de A y escribir $x \in A_\alpha$.

El conjunto borroso A es así la unión de los diferentes niveles de pertenencia A_α . La representación gráfica de A se asemeja a un mapa con diferentes curvas de nivel.

Relaciones borrosas

Sean A y B dos subconjuntos borrosos de X e Y, se define el *producto cartesiano* de dos subconjuntos borrosos

$$A \times B (x,y) = \min \{A(x), B(y)\}.$$

De forma similar al modelo clásico, las relaciones borrosas se definen como subconjuntos (borrosos) del producto cartesiano.

Cuando $A = B$ se dice que R es una relación borrosa en A.

Aunque A y B sean nítidos, se pueden definir entre ellos relaciones de tipo borroso; y éstas serán las que utilizaremos con mayor frecuencia en nuestro estudio sobre la medida escolar. En este caso la relación viene definida por una matriz simétrica y podemos establecer una propiedad de gran importancia:

R es una relación binaria borrosa entre A y B si y sólo si para todo $\alpha \in [0,1]$ se verifica que R_α es una relación nítida entre A_α y B_α .

Esta propiedad nos permite una representación de las relaciones borrosas con «curvas de nivel».

Ejemplo:

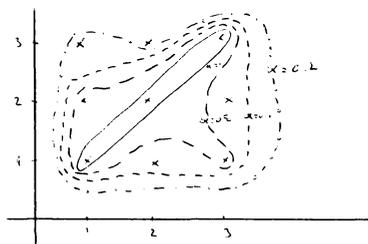
Sea $A = \{1,2,3\}$ y sea

$$[R] = \begin{bmatrix} 1,0 & 0,8 & 0,2 \\ 0,4 & 1,0 & 0 \\ 0,8 & 0,4 & 1,0 \end{bmatrix}$$

los niveles de nitidez de R serán:

$$\begin{aligned} R_1 &= \{(1,1), (2,2), (3,3)\}, \\ R_{0,8} &= R_1 \vee \{(1,2), (3,1)\}, \\ R_{0,4} &= R_{0,8} \vee \{(2,1), (3,2)\}, \\ R_{0,2} &= R_{0,4} \vee \{(1,3)\}. \end{aligned}$$

y la representación en curvas de nivel de la relación tendrá el siguiente aspecto:



Las conocidas propiedades reflexiva, simétrica, antisimétrica y transitiva para las relaciones borrosas se pueden definir (a través de los niveles de nitidez) con una formulación similar a la que se utilizaba en las clásicas.

Del mismo modo que en la teoría clásica de conjuntos ocupan un lugar fundamental las relaciones de equivalencia y las relaciones de orden, en la teoría de los conjuntos borrosos son muy importantes los siguientes tipos de relaciones:

Semejanzas: cuando son reflexivas y simétricas. También se les llama proximidades.

Preórdenes: si poseen las propiedades reflexiva y transitiva.

Equivalencias borrosas: son los preórdenes simétricos (o semejanzas transitivas). También son conocidas como similitudes.

Órdenes borrosos: son los preórdenes antisimétricos.

Consideraciones sobre la TCB

Como afirma Llorenç Valverde (1984): «La teoría de los conjuntos difusos, entendida en términos de su objetivo práctico esencial, pretende facilitar el manejo de predicados y relaciones vagas que necesariamente aparecen con el intento de entender y controlar sistemas complejos, especialmente sistemas en los que se debe operar en ambiente impreciso o incierto y para los que las técnicas estadísticas no tienen siempre respuesta».

Con los conjuntos borrosos podemos evitar la introducción de una precisión artificial en conceptos que son intrínsecamente vagos. La teoría de Zadeh permite la construcción de modelos para la descripción de conjuntos y relaciones que, en la teoría clásica, eran considerados como «mal definidos» pero que están presentes en la realidad y en los razonamientos cotidianos.

Asegura Trillas (1980): «el éxito de la teoría se ha producido en momentos en que la economía, la sociología, la lingüística, la psicología, ... requieren un nuevo nivel de matematización para seguir avanzando». Nosotros hemos de añadir que esta misma necesidad se presenta también (y cada día con mayor insistencia) en las Ciencias de la Educación.

2. UN MODELO BORROSO PARA LA DESCRIPCIÓN DE LA MEDICIÓN ESCOLAR

En cualquier ámbito científico una medición en un conjunto A es una aplicación

$$m: A \longrightarrow R$$

que asigna a cada elemento de A (denominados elementos u objetos medibles) un valor $v \in R$ y uno sólo, y una serie de axiomas que definen las reglas de la medida

y que pueden variar según las características y necesidades del modelo métrico.

Estos axiomas pueden concretarse en forma de instrumento con unas reglas específicas de medida. En general toda medición en el ámbito escolar se realiza utilizando un instrumento que básicamente consta de una entrada de información y de unos criterios que permiten cuantificar la respuesta.

Para caracterizar las respuestas hemos de distinguir si las puntuaciones son dicotómicas o continuas. Para el primer caso toda respuesta a un test I se puede caracterizar como el subconjunto nítido de los ítems correctos. Pero si las puntuaciones son continuas (en el intervalo $[0,1]$) toda respuesta vendrá caracterizada por el subconjunto borroso definido por la función de pertenencia que asigna a cada ítem su puntuación.

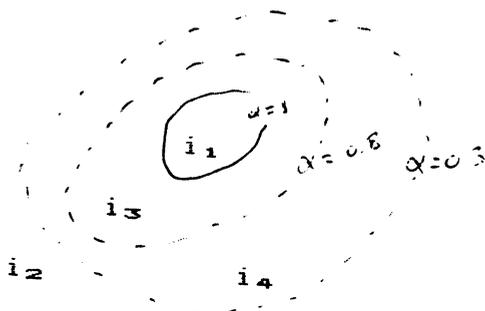
Ejemplo: Sea $I = \{i_1, i_2, i_3, i_4\}$
 un test con cuatro ítems;
 sea y la respuesta definida por la aplicación

$P_y: I$	\longrightarrow	$[0,1]$
i_1	\longrightarrow	1
i_2	\longrightarrow	0
i_3	\longrightarrow	0,8
i_4	\longrightarrow	0,3

En este caso el alumno ha dado una respuesta correcta al ítem 1 y ha realizado incorrectamente el ítem 2. Pero el resto de ítems no están «ni bien, ni mal»: tienen un cierto grado de corrección. El subconjunto de los ítems correctos de I en la respuesta dada por este alumno será

$$\{i_1/1, i_2/0, i_3/0,8, i_4/0,3\}$$

que es borroso y que viene definido por la aplicación p_y con diferentes niveles de pertenencia.



Con esta descripción de las respuestas a través de la TCB, podemos definir la inclusión, la unión y la intersección de respuestas, así como la respuesta nula y la total. Y queda abierto el camino para una formulación axiomática de la medida escolar (ver Gil Quesada, 1990).

3. EL MODELO DE LA ZONA DE CORTE BORROSO

El modelo clásico del punto de corte (característico de las mediciones criterioles) tiene como función el establecimiento de un standard de suficiencia. Para ello se toma un $k \in [0,1]$ fijo, denominado punto de corte (o punto de separación standard) y se decide que el nivel de aprendizaje de un alumno es

- suficiente si su puntuación p verifica $p \geq k$.
- insuficiente si su puntuación p verifica $p < k$.

Es decir, para todo punto de corte k se define una aplicación $F(p) = 0$ si $p < k$ y $F(p) = 1$ si $p \geq k$ que dicotomiza la variable continua y produce una reducción de la información que nos proporcionan los mecanismos de medida. Este paso reductor conlleva una contradicción lógica importante que permite trasladar al ámbito de la medida escolar la conocida paradoja «todos los hombres son altos» que puede situarse en la base de la Teoría de los Conjuntos Borrosos: Si se toma 5 como puntuación de corte, pero consideramos que nuestro instrumento de medida no aprecia más allá de las centésimas, deberíamos aceptar una puntuación de 4.999 como apta. Y una vez aceptado esto: de milésima en milésima, llegaríamos a la conclusión de que toda puntuación debe ser aceptada como suficiente.

Esta paradoja está en relación con la complejidad de la realidad escolar y con las limitaciones de los instrumentos de medida de que disponemos. El establecimiento del punto de corte trata de evitarlas, pero para ello ha de forzar la realidad borrosa mediante un modelo nítido.

La separación de paso entre las puntuaciones no parece que sea tan nítida. Al menos es frecuente que en la práctica escolar se acepte la existencia de una zona en la que se sitúan puntuaciones que corresponden a los alumnos cuyo grado de suficiencia es dudoso.

La Teoría de Conjuntos clásica parte de una concepción dicotómica de la pertenencia, ligada con el principio del tercero excluido de la lógica bivalente. En este marco un alumno es apto o no lo es. Y no cabe ninguna otra posibilidad.

El modelo apropiado ha de describir la realidad de la medición contemplando toda su complejidad: las puntuaciones nítidamente insuficientes, las dudosas y las suficientes nítidas. Y éste es el objetivo del modelo que a continuación exponemos:

Dado el conjunto Y de respuestas, que hemos identificado con los subconjuntos borrosos del examen I , podemos considerar que todos los valores del intervalo $[0,1]$ son puntuaciones posibles.

Si consideráramos que las puntuaciones suficientes son un subconjunto nítido de $[0,1]$, con punto de corte k , bastaría con definir una función característica F

$$\begin{array}{l} F: [0,1] \longrightarrow \{0,1\} \\ p \longrightarrow F(p) \end{array}$$

(siendo $F(p) = 0$ si $p < k$ y $F(p) = 1$ si $p \geq k$) que correspondería al modelo de punto de corte nítido.

Pero si consideramos que el concepto «aprobado» no es nítido tendremos que caracterizarlo utilizando la TCB: el subconjunto A de las puntuaciones aprobadas se definirá mediante una función de pertenencia

$$\begin{array}{l} F': [0,1] \longrightarrow [0,1] \\ p \longrightarrow F'(p) \end{array}$$

siendo $F'(p) = 0$ si $p \leq \alpha$, $F'(p) = 1$ si $p \geq \beta$ y $F'(p) = f'(p) \in (0,1)$, si $p \in (\alpha, \beta)$

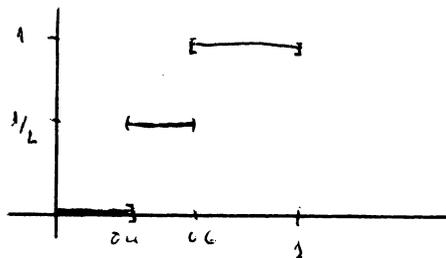
Las puntuaciones del subconjunto borroso A mayores que β tendrán un grado de pertenencia 1 (son aprobados claros) y las menores que α tendrán un grado de pertenencia 0; las puntuaciones intermedias pertenecen a A con un grado de pertenencia variable definido por el examinador (ya sea por una fórmula algebraica o por cualquier otro método bien determinado).

A este modelo descriptivo lo denominamos *Modelo de la zona de corte borroso*, y lo presentamos como una alternativa (y también como una generalización) conceptual al usual punto de corte característicos de los modelos criteriosales en la escuela.

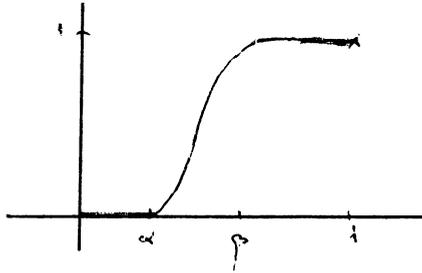
Para clarificar el modelo veamos la representación gráfica de algunos ejemplos:

- 1) Si $f'(p) = 1/2$, $\alpha = 0,4$ y $\beta = 0,5$ estaríamos dando como
 - aprobados claros (nivel de nitidez 1) a los alumnos que obtienen una puntuación superior a 0,5,
 - aprobados (o suspensos) borrosos con nivel 1/2 a los que obtienen una puntuación entre 0,4 y 0,5, y
 - suspensos claros a los que tienen $p < 0,4$.

Y la función F' se representaría



2) Obtendríamos un modelo más gradual si tomamos como función de pertenencia de la zona de corte borroso a la función S de Zadeh cuya representación es como sigue:



y corresponde a una f' definida por

$$f'(p) = 2[(p-\alpha)/(\beta-\alpha)]^2, \text{ si } \alpha \leq p \leq t$$

$$f'(p) = 1-2[(p-\beta)/(\alpha-\beta)]^2, \text{ si } t \leq p \leq \beta.$$

4. LAS RELACIONES BORROSAS EN EL ANÁLISIS DE LA MEDIDA ESCOLAR

Como ejemplo de las posibilidades que ofrece la utilización de las relaciones borrosas en el análisis de la medida, presentamos dos modelos para el estudio de la estructura de un test. El primero utiliza una relación de proximidad que cuantifica hasta que punto son homogéneas las respuestas que el grupo de alumnos da a los distintos ítems del test. El segundo, mediante un preorden borroso, ofrece la posibilidad de adentrarse en la naturaleza jerárquica del test.

4.1. Análisis de la estructura homogénea

Para el estudio de la homogeneidad de un test hemos de recurrir a algún índice que cuantifique el grado de semejanza que se puede apreciar en las respuestas a los distintos ítems del test. Son numerosos los instrumentos existentes en el ámbito de la Estadística que pueden cumplir esta función. Como ejemplo (y por motivos teóricos que la brevedad de este artículo no nos permite desarrollar) hemos elegido como punto de partida el índice p_0 de Hambleton (Berk, 1980).

Dados dos elementos de medida e_1, e_2 (ítems o test) se define $p_0(e_1, e_2)$ como la proporción de alumnos que responden correctamente a los dos elementos más la proporción de los que responden incorrectamente (también a los dos): $p_{00} + p_{11}$.

Generalizando este índice a puntuaciones no necesariamente dicotómicas y centrándonos en dos ítems i, i' tendríamos

— para dos puntuaciones p y p' de un alumno a_k en i e i'

$$c_k(i, i') = 1 - |F(p) - F(p')|$$

siendo F la función de pertenencia de la zona de corte borroso (si se toma un modelo de punto de corte nítido: la F simplemente dicotomiza).

— para todo el grupo A de alumnos, la concordancia entre dos ítems vendría definida por $c(i,i') = 1/n \sum_k c_k(i,i')$.

La concordancia entre dos ítems cualesquiera de un test I establece en este conjunto una relación borrosa. El grado de relación entre cada par de ítems permite definir una matriz de interconcordancia simétrica, con 1 en la diagonal principal y valores entre 0 y 1 en el resto.

Esta relación posee las propiedades reflexiva y simétrica, ya que éstas se verifican en las relaciones nítidas definidas para cada uno de los niveles α . En consecuencia, se trata de una proximidad o semejanza; pero no es una relación de equivalencia. Pues, aunque en casos específicos puede darse también la transitividad, en general esta propiedad no se cumple para la relación de concordancia. Se puede demostrar que la concordancia es un operador de indistinguibilidad, concepto de gran importancia en la TCB pero en el que no vamos a entrar en este artículo.

No podemos establecer una clasificación en el conjunto de ítems, pero si podemos situar los elementos del test en matrices y grafos que pongan de manifiesto su semejanza. Estos elementos analíticos se podrían situar entre las diversas técnicas del «cluster analysis» u otras propias de los métodos multivariados.

Grafos de proximidad

El proceso que se ha de seguir para el estudio de los grafos de proximidad entre los ítems consta de los siguientes pasos:

1. Se parte de la tabla de puntuaciones obtenidas por el grupo de alumnos en los distintos ítems del test.
2. El cálculo de la concordancia entre los distintos ítems permite la construcción de la matriz de la relación borrosa.
3. Para cada nivel de concordancia obtendremos una matriz de ceros o unos.
4. Asociado a cada una de estas matrices obtendremos un grafo de proximidad en el conjunto de ítems.

Ejemplo: Consideremos las siguientes puntuaciones obtenidas por diez alumnos en un examen de cinco ítems:

	i1	i2	i3	i4	i5
a1	1	0	0,3	0,8	0,8
a2	0,1	0	0	0	0
a3	1	0,8	0,2	1	0,8
a4	1	1	0	0	0
a5	0,2	0	0	0	0
a6	1	1	0	0	0,2
a7	1	0,7	0,3	0,9	0,8
a8	0	0	0	0	0
a9	1	0	0	0	1
a10	1	0,8	0	0,9	1

Aunque en el estudio anterior hemos definido la concordancia para la zona de corte borroso, en este ejemplo tomaremos un modelo de corte nítido ($k = 0,5$) para facilitar la comprensión de lo que aquí es esencial: los grafos de proximidad o semejanza.

Para esta restricción a $k = 0,5$ y para la tabla de puntuaciones de la que partimos la matriz que define la relación borrosa de concordancia inter-ítems será:

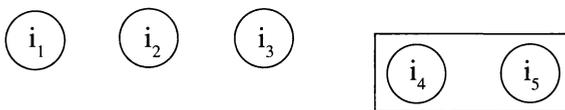
$$\begin{bmatrix} 1 & 0,8 & 0,4 & 0,8 & 0,8 \\ 0,8 & 1 & 0,6 & 0,6 & 0,6 \\ 0,4 & 0,6 & 1 & 0,6 & 0,6 \\ 0,8 & 0,6 & 0,6 & 1 & 1 \\ 0,8 & 0,6 & 0,6 & 1 & 1 \end{bmatrix}$$

Pasando al estudio de los distintos niveles de concordancia inter-ítems, tenemos que

1. La relación nítida para $\alpha=1$ sería

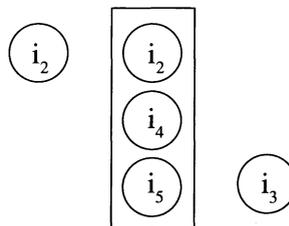
$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} = R_{\alpha=1}$$

que definiría el siguiente gráfico de proximidad:



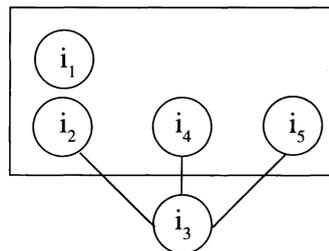
2. Para $\alpha=0,8$ la matriz y el gráfico de proximidad serían:

$$\begin{bmatrix} 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 \end{bmatrix} = R_{cr=0,8}$$



3. Y para $\alpha=0,6$ se tendría

$$\begin{bmatrix} 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} = R_{cr=0,6}$$



Índice de homogeneidad del test

A partir de la matriz de concordancia entre los ítems, se puede introducir un Índice de Homogeneidad del test definido por

$$H = \frac{\sum \sum c(i,j) - n}{n^2 - n}$$

que no es más que la medida aritmética de los valores de la matriz de concordancia inter-ítems que no están en la diagonal.

Para el ejemplo anterior sería:

$$H = \frac{2,8 + 2,6 + 2,2 + 3 + 3}{25 - 5} = \frac{13,6}{20} = 0,68$$

Este Índice de Homogeneidad H del test es una cuantificación del grado de concordancia global existente entre todos los ítems que componen el test.

5.2. Análisis de la estructura jerárquica del test

La homogeneidad de los ítems que componen un test es una de las características más valoradas en cuanto a la estructura de un test. Y en ciertos aspectos se podría llegar a decir que ha sido el núcleo exclusivo de todo análisis estructural. La fiabilidad podría considerarse como una exigencia de homogeneidad.

Esta excesiva valoración de la homogeneidad tal vez esté justificado en psicometría; sin embargo en las Ciencias de la Educación, esta exigencia debería relativizarse. De hecho varios autores, entre ellos Berk (1980), ya se han planteado la necesidad de la ampliación del análisis a otras propiedades.

Si un test pretende medir el aprendizaje que un cierto grupo de alumnos ha realizado, al ser sometidos a un determinado programa de enseñanza, interesará saber no sólo si los alumnos son suficientemente expertos sino también el nivel al que han llegado en su aprendizaje. Ítems con diferentes grados de dificultad pueden aportar información valiosa en este aspecto. Aparece por tanto la necesidad de estructuras jerarquizadas.

Por otro lado, un test construido con ítems jerarquizados posibilita un análisis cualitativo y no sólo cuantitativo de las respuestas de los alumnos. Si la realización de un ítem está condicionado a la de otro, las respuesta a estos dos ítems aportan una información cualitativamente distinta de la que se pudiera extraer de dos ítems independientes y con semejante grado de dificultad. Con la jerarquización se puede analizar propiedades topológicas (forma, lagunas, respuestas inusuales, etc.) que enriquecen la medición.

Pero la relación de jerarquización no es nítida; es necesario establecer un modelo de relación difusa que describa el grado de subordinación de un ítem en otro.

Índice de jerarquización para dos ítems (puntuación dicotómica)

Sean i, i' dos ítems del test I. Sea A un grupo de n alumnos a los que se ha administrado el test I. Definimos el índice de jerarquización $J(i, i')$ del ítem i en el i' :

$$J(i, i') = \max \{j(i, i'), 0\}, \text{ siendo}$$

$$j(i, i') = \frac{n(i \cap \bar{i}') - n(\bar{i} \cap i')}{n(i \cap \bar{i}') + n(\bar{i} \cap i')}$$

donde $n(i \cap i')$ indica el número de alumnos que han obtenido 1 en i y 0 en i' . Y análogamente para los otros términos.

Cuando el denominador es cero, definimos $j(i, i') = 0$.

Cuando dos ítems tienen una respuesta idéntica por el grupo de alumnos se verifica que tanto el numerador como el denominador son nulos. Si consideramos una jerarquización estricta hemos de definir el valor de j como 0; pero si preferimos una jerarquización no estricta (similar al \leq en \mathbb{R}) podemos asignarle el valor de 1.

Si $j(i, i') > 0$, diremos que i está *subordinado* a i' .

Si i está totalmente subordinado a i' , de tal modo que

- todos los alumnos que hacen bien i' han hecho correctamente i y además
 - algún alumno responde correctamente i y no i' ,
- se verifica

$$J(i,i') = j(i,i') = 1.$$

En general los valores de $J(i,i')$ tendrán un rango $[0,1]$.

Cuando $j(i,i') < 0$ el valor de $J(i,i') = 0$ y esto será indicativo de que i' está subordinado a i ; es decir

$$J(i',i) = j(i',i) > 0$$

Ejemplos:

a)

i'	0	0	0	0	0	1	1	1	1	1
i	1	1	1	1	1	1	1	1	0	0

$$J(i,i') = 3/7 = 0,4$$

c)

i'	0	0	0	1	1	1
i	1	1	1	0	0	0

$$J(i,i') = 0/6 = 0$$

d)

i'	0	0	0	0	0	1	1	1	1	1
i	1	1	1	1	1	1	1	1	0	0

$$J(i,i') = 3/7 = 0,4$$

e)

i'	0	0	0	1	1	1	1	1	0	0
i	1	1	1	1	1	1	1	1	0	0

$$J(i,i') = 3/3 = 1$$

Se puede demostrar que para que exista algún grado de subordinación de i en i' es necesario y suficiente que el ítem i sea más fácil que i' . No obstante la diferencia entre los índices de facilidad de dos ítems no determina el grado de jerarquización entre ellos (ver los ejemplos d y e anteriores).

La jerarquización en el test

El estudio que anteriormente hemos realizado para un par de ítems se puede extender a todo el test I . La jerarquización J define una relación difusa en I que se puede representar por una matriz cuadrada de k filas y columnas (k = número de ítems).

Esta relación difusa es transitiva ya que

$$J(i,i'') > \min \{J(i,i'), J(i',i'')\}.$$

En otras palabras si, para un nivel α , i está subordinado en i' , e i' lo está en i'' , se verifica que i está subordinado en i'' .

Si consideramos una jerarquización no estricta: la relación de jerarquización, además de transitiva, es también reflexiva. Tendremos pues un preorden borroso.

Jerarquización en modelos no dicotómicos

El modelo de análisis de la jerarquización para puntuaciones dicotómicas que hemos descrito en los puntos anteriores se puede generalizar para el caso de calificaciones en [0,1] y con un modelo de corte borroso (siendo F la función de pertenencia). La fórmula para el caso más general sería

$$j(i,i') = [\sum(F(p_a(i)) - F(p_a(i')))] : [\sum|F(p_a(i)) - F(p_a(i'))|]$$

Grafos asociados a la jerarquización

A partir de la matriz de la relación difusa J(i,i') podemos obtener tantas relaciones nítidas como valores distintos existan en la matriz. Estos valores definirán los distintos niveles de nitidez de cada una de las relaciones y podremos asociarles grafos de flechas que representen la jerarquización de los ítems del test I. Teniendo en cuenta la transitividad de J, no es necesario ni conveniente indicar la flecha de i a i'' si ya se ha indicado la de i a i' y la de i' a i''.

Cada uno de los grafos tendrá un número de flechas r que denominaremos *peso del grafo* y que indicará el grado de complejidad de la relación nítida correspondiente. Sea R = max r. El grafo que corresponde a este valor R es el que presenta mayor complejidad y el que nos puede aportar mayor información.

Ejemplo de aplicación:

Consideremos la tabla de puntuaciones

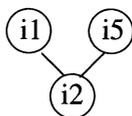
	i1	i2	i3	i4	i5
a1	1	1	1	0	0
a2	0	0	1	0	0
a3	1	1	1	0	0
a4	0	0	0	0	0
a5	1	1	1	1	1
a6	0	1	0	1	0
a7	0	0	0	1	0
a8	1	1	0	0	0
a9	0	1	1	1	1
a10	1	1	0	0	1

La matriz de jerarquización sería

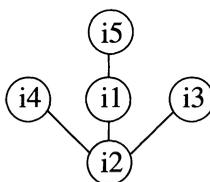
	i1	i2	i3	i4	i5
i1	0	0	0	0,14	0,50
i2	1	0	0,50	0,75	1
i3	0	0	0	0,20	0,50
i4	0	0	0	0	0,33
i5	0	0	0	0	0

que nos permite definir relaciones nítidas para los distintos niveles de la relación difusa.

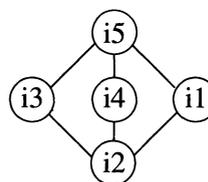
$\alpha = 1$



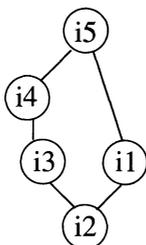
$\alpha = 0,5$



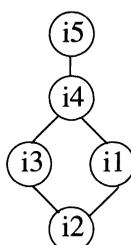
$\alpha = 0,33$



$\alpha = 0,2$



$\alpha = 0,14$

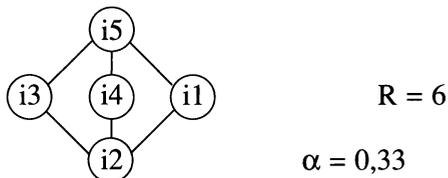


El gráfico que corresponde al nivel 0,33 tiene el mayor peso $r = 6$ y, por tanto será el que nos aporte una mayor información.

Aplicabilidad al análisis de ítems

En el ejemplo anterior hemos visto como, a partir de la matriz de jerarquización que establece la relación borrosa, podemos considerar los distintos niveles de nitidez que definen estructuras jerárquicas en I (relaciones nítidas). La cantidad y calidad de la información de cada una de ellas será variable y podemos escoger el nivel de mayor peso R que es el que aporta mayor riqueza de análisis.

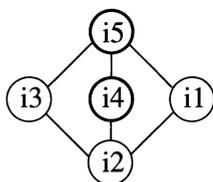
En nuestro ejemplo el grafo a considerar era el de nivel de nitidez 0,33.



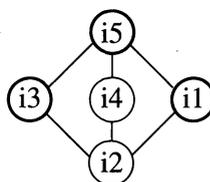
A partir de este grafo podemos abordar el estudio de las respuestas de cada uno de los alumnos desde una perspectiva de una mayor riqueza, ya que posibilita detectar las respuestas inusuales y analizar los errores en los niveles más altos a partir de los cometidos en los niveles inferiores.

Siguiendo en el esquema del ejemplo considerado, veamos algunos casos que permiten este estudio topológico:

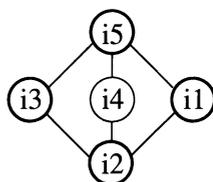
alumno a1:



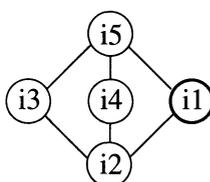
alumno a6:



alumno a7:



alumno a9:



- En el alumno a1 el fallo en el i5 puede deberse a la no realización de i4.
- En el alumno a6 serían los errores en i3 y i1 los que condicionarían el error en i5.
- Los alumnos a7 y a9 presentan una respuesta claramente inusual.

Para un estudio más profundo del análisis de las respuestas inusuales remitimos al lector al sugestivo estudio propuesto por Sato, T. (1980) y a la complementación criterial dada en Gil Quesada, X. (1990).

BIBLIOGRAFÍA

- BERK, R. A. (Ed.) (1980a): *Criterion Referenced Measurement: The State of the Art*. Baltimore: Johns Hopkins Univ. Press.
- GIL QUESADA, X. (1990): *La Teoría de los Conjuntos Borrosos en la Medición Escolar*. Tesis Doctoral.
- HAMBLETON, R. K. (1980b): Test score validity and standard-setting methods. En R. A. BERK (Ed.): *Criterion-referenced Measurement: The state of the art*. Baltimore, MD: Johns Hopkins Univ. Press, 80-123.
- HAMBLETON, R. K. y EIGNOR, D. R. (1978b): Adaptive testing applied to hierarchical-structured objectives-based curricula. En D. J. WEISS (ed.): *Proceedings of the Second Computerized Adaptive Testing Conference*. Minneapolis: Computerized Adaptive Testing Laboratory, Univ. of Minnesota, 290-311.
- HAMBLETON, R. K.; SWAMINATHAN, H.; ALGINA, J. y COULSON, D. B. (1978): «Criterion-referenced testing and measurement: A review of Technical Issues and Developments». *Review of Educational Research*, vol. 48(1), 1-47.
- RUSPINI, E. (1982): «Recent Developments in Fuzzy Clustering». En R. R. YAGER (Ed.): *Fuzzy Set and Possibility Theory: Recent Development*. New York: Pergamon Press, 133-147.
- SATO, T. (1980): *The S-P Chart and the caution index*. Tokyo, Japan: Computer and Communications Systems Research Laboratories, Nippon Electric Company.
- SMITTHSON, M. (1987): *Fuzzy Set Analysis for Behavioral and Social Sciences*. Springer-Verlag New York Inc.
- SWAMINATHAN, H.; HAMBLETON, R. K. y ALGINA, J. (1974): «Reliability of criterion-referenced tests: A decision-theoretic formulation». *Journal of Educational Measurement*.
- TRILLAS, E. (1980): *Conjuntos Borrosos*. Barcelona: Vicens-Vives.
- VALVERDE, L. (1984): *Lógica Polivalente y Vaguedad Lingüística*. Barcelona: ETSA, UPB.
- VALVERDE, L. (1985): «On the structure of F-indistinguishability operators». *Fuzzy Sets and Systems*, 17, 313-328.
- ZADEH, L. A. (1965): «Fuzzy Sets». *Information and Control*, 8(3), 338-353.
- ZADEH, L. A. (1965): «Fuzzy Sets and Systems». En J. FOX (Ed.): *System Theory*. New York: Polytechnic Press, 29-37.
- ZADEH, L. A. (1971): «Similarity Relations and Fuzzy Orderings». *Information Sciences*, 3, 177-200.
- ZADEH, L. A. (1977): *Fuzzy Sets and Their Applications to Pattern Classifications and Cluster Analysis*. En J. van RYZIN (Ed.): *Classification and Clustering*. New York: Academic Press.
- ZADEH, L. A. (1983c): «Linguistic variables, approximate reasoning and dispositions». *Med. Inform.*, 8, 173-186.
- ZADEH, L. A. (1984): «A theory of commonsense knowledge». En H. K. SKALA et al. (Eds.): *Aspects of Vagueness*, Reidel, Dordrecht, 257-295.

ESTUDIO DE LAS CALIFICACIONES ESCOLARES MEDIANTE ANÁLISIS EXPLORATORIO DE DATOS

por

*Lluís Salafranca Cosialls, Montserrat Freixa Blanxart
y Fco. Javier Ormazábal Unzué*

RESUMEN

Tukey 1977 en su libro «Exploratory Data Analysis» expone un conjunto de nuevas técnicas estadísticas resistentes y robustas que intentan descubrir patrones o modelos, anomalías o errores en los datos, para ello se valen de importantes innovaciones principalmente gráficas. Estas técnicas no sólo constituyen un complemento a las clásicas sino también una valiosa alternativa.

En este artículo, se evidencian las ventajas de aplicar algunas de estas técnicas a las calificaciones escolares.

ABSTRACT

Tukey in his book «Exploratory Data Analysis» exposed in 1977 an ensemble of new statistical techniques, both robust and resistant, intending to find data models and outlines, errors and anomalies. With this purpose new devices are introduced most of they graphical. These are not only complementary to the classical ones but also they are valuable alternatives.

In this paper are shown the promising possibilities of the use of such technical procedures in school's problems.

INTRODUCCIÓN

Tukey en su libro «Exploratory Data Analysis» (1977) E.D.A., desarrolla una serie de nuevas técnicas gráficas y analíticas para conseguir un conocimiento previo de los datos a analizar, siempre desde una perspectiva exploratoria.

El análisis exploratorio de datos propugna un cambio de actitud y de enfoque metodológico ante el análisis de datos.

Postula que es necesario *explorar* detenidamente los datos antes de empezar cualquier análisis.

La Estadística Descriptiva clásica se ocupa en recoger, ordenar y representar los datos, normalmente en forma de tablas y agrupando los datos en intervalos para representarlos gráficamente. Calcula principalmente estadísticos basados en la distancia, con datos medidos en escala de intervalo y toma como índice de referencia la media.

El E.D.A. tiene los mismos objetivos, pero pretende además *detectar anomalías o errores* en las distribuciones univariantes de los datos de forma que éstos no incidan o invaliden posteriores análisis. También intenta *descubrir en los datos patrones o modelos*. Para ello incorpora nuevas técnicas gráficas y busca estadísticos resistentes y robustos basados principalmente en el orden y tomando como referencia la mediana.

Su sencillez y rapidez de cálculo la hacen sumamente útil en Ciencias Sociales, Humanas y de la Salud para explorar distribuciones univariantes, así como estructuras de relación entre variables.

El E.D.A. tiene cinco características principales:

- 1) *Sus representaciones gráficas* nos revelan visualmente el comportamiento de los datos y la estructura de conjunto.
- 2) Se sirve de *la transformación de los datos*, que consiste en encontrar la escala que más simplifique y clarifique el análisis, como por ejemplo con el uso de funciones matemáticas simples como raíz cuadrada, logaritmos, etc.
- 3) Valora *la resistencia*, propiedad que presentan ciertos estadísticos de ser poco sensibles a la influencia de unos valores muy distantes de la mayoría de los de la distribución.
- 4) Busca estadísticos *robustos*, propiedad que presentan algunos estadísticos que les hace poco sensibles a desviaciones de los supuestos básicos.
- 5) Pone mucha atención en el análisis de *residuales*, es decir en las diferencias que hay entre los datos reales y el resultado de un ajuste exploratorio a un modelo previamente determinado o subyacente.

Cuando buscamos relaciones entre las variables, el E.D.A. es especialmente adecuado en las disciplinas anteriormente mencionadas donde los modelos sustantivos son complejos y las variables han sido medidas en todo tipo de escalas, nominal, ordinal, de intervalo y de razón y los datos están sujetos a gran variabilidad.

Así, dado el desconocimiento de los verdaderos modelos y teorías que generalmente subyacen en estos campos, los análisis mediante E.D.A. ayudan a descubrir tendencias, patrones de conducta, conductas diferenciales, formación de actitudes y evaluación del cambio.

Cabe destacar así mismo que las técnicas E.D.A. no sólo constituyen un complemento a las técnicas estadísticas clásicas si no también una valiosa alternativa en caso de incumplimiento de alguna condición de aplicación, puesto que no son tan restrictivas en sus supuestos.

En realidad el investigador necesita usar las técnicas estadísticas exploratorias y confirmatorias. Las técnicas exploratorias ayudan a comprobar las condiciones de aplicación de las pruebas de hipótesis, a detectar errores o valores anómalos, a buscar la mejor transformación cuando es necesaria, etc. En general dan una visión distinta, previa, pero complementaria a la confirmatoria. Todo ello repercute en una mejor calidad del análisis de datos globalmente entendido.

DIAGRAMA DE TALLO Y HOJAS

Tukey (1977) idea una representación gráfica, llamada diagrama de «tallo y hojas» (Steam and Leaf), para variables cuantitativas.

El diagrama de tallo y hojas es un procedimiento semi-gráfico de presentar la información de variables cuantitativas. Es en realidad, una representación visual de la distribución de una variable, que intenta respetar su información cuantitativa al máximo.

Para construir el diagrama de tallo y hojas se hace una tabla con dos columnas separados por una línea y cada dato se desglosa en varias cifras.

Para datos con dos dígitos, como por ejemplo 29, se escribe a la izquierda de la línea los dígitos de las decenas —que forman el tallo— y a la derecha las unidades serán las hojas.

Ejemplo: 29 se representa

2 9
tallo hoja

Para datos con tres dígitos el tallo puede estar formado por los dígitos de las centenas y decenas que se escribirán a la izquierda, separadas de las unidades.

Ejemplo: 567 se representa

56 7
tallo hoja

Lo primero que hay que hacer es buscar la unidad que es la hoja. Cada tallo define una clase y se escribe una sola vez. El número de hojas representa la frecuencia de dicha clase.

Ejemplo 54, 57, 69, 67, 43, 59, 62, 36, 74, 67, 46, 75,

4	36	4	3	representa 43
5	479			
6	2.779	n=12		
7	45			
tallo	hojas			

El número de hojas es igual al número de casos.

En Estadística Descriptiva clásica el histograma es una representación gráfica muy frecuente en la que toda la información se concentra en unos intervalos.

El diagrama de tallo y hojas se parece a un histograma que retiene y ordena todos los datos, aunque estén repetidos, no perdiendo ninguno de ellos y nos da una buena representación de la forma de la distribución.

Las ventajas del Diagrama de tallo y hojas sobre el histograma se pueden concretar en tres niveles.

- 1) Es más fácil de preparar manualmente.
- 2) Puede examinarse más detalladamente que el histograma por que las barras de un histograma pueden ocultar puntos dentro de ellas pero esto no ocurre con el diagrama de tallo y hojas ya que este retiene los valores numéricos de los datos.
- 3) El diagrama de tallo y hojas es un gráfico muy flexible, que permite varios tipos, mientras que el histograma es un gráfico más rígido.

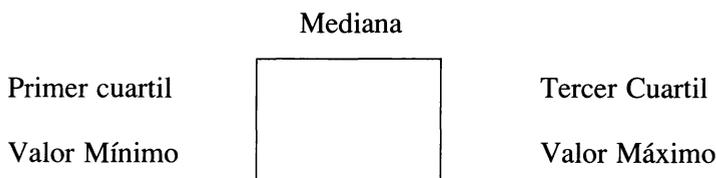
Existen diferentes representaciones del diagrama de «tallo y hojas», a base de diferentes subdivisiones o cambios de unidad. Véase B. Erickson y T. Nosanchuk. (1983).

Hoaglin, Mosteller y Tukey (1983) y Velleman y Hoaglin (1981) coinciden en que el diagrama de tallo y hojas presenta las siguientes características que lo convierten en un gráfico muy útil para «ver» en los datos las siguientes características:

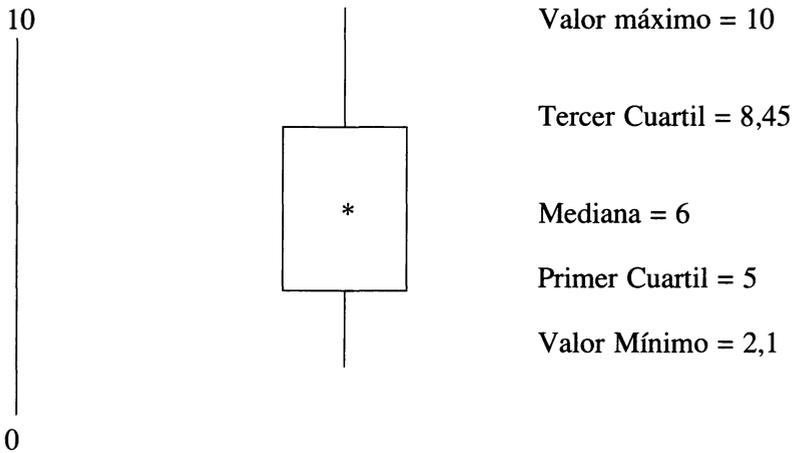
- a) La dispersión de los valores.
- b) En dónde están concentrados los valores.
- c) La simetría de la serie.
- d) Si se presentan agujeros (vacíos en donde no se encuentran valores).

Diagrama de Caja

Tukey (1977) presenta una nueva representación gráfica que denomina «diagrama de caja» (Box Plot). El esqueleto del más simple de los «diagramas de caja» se construye a partir de la Mediana, los cuartiles y los valores máximo y mínimo.



Dentro de la «caja» se encuentra el 50% central de valores de la distribución. Se suele disponer de la siguiente manera:



Sin embargo normalmente el diagrama de caja se hace más completo ya que detecta los valores alejados (outliers) y estudia hasta que punto se alejan de la normalidad.

El diagrama de caja nos muestra la estructura de la serie de datos. A partir de esta representación gráfica se puede ver claramente, la dispersión, la simetría, el aspecto y el alcance de las colas y los valores alejados de una distribución, así como la localización de un valor determinado.

El diagrama de caja resulta especialmente útil para comparar varias distribuciones a la vez. Dibujando en paralelo el diagrama de caja de cada distribución, podemos disponer de una impresión visual rápida de las similitudes y diferencias entre las distribuciones, así como de las características de las mismas. El diagrama de caja sugiere con frecuencia la mejor transformación de los datos cuando ésta sea necesaria.

Ejemplo aplicado a las calificaciones escolares

Es ya práctica habitual en muchos colegios que en el boletín de calificaciones después de cada evaluación, además de las notas y actitud en cada asignatura, alguna información estadística, dirigida a los padres, que pretende comparar la calificación del alumno con la media del grupo.

Normalmente la información es como sigue

	Estudio compartivo						
	media	def.	insuf.	suf.	bueno	not.	sob.
cal. alumno	3,5	*****					
med. grupo	5,46	*****					

En esta información se observa la diferencia entre la nota media del grupo y la calificación del alumno.

No obstante dado que no se acostumbra a dar ninguna medida de dispersión es imposible ver la posición del alumno dentro del grupo.

Además en este estudio comparativo se utiliza normalmente la media como índice descriptivo. La media es un estadístico no resistente ya que puede estar afectada por las notas altas o bajas de dos o tres sujetos no representativos del grupo.

Estudiemos dos clases con las puntuaciones y los estadísticos que se detallan en la tabla 1. Las dos clases tienen prácticamente la misma media (media clase 1 = 5,47 media clase 2 = 5,46). Entonces dos alumnos, uno de la clase 1 y otro de la clase 2 con una puntuación de 3,5, por ejemplo, tendrían una gráfica de estudio comparativo igual en las dos clases. No obstante si analizamos más detenidamente las distribuciones veremos que tienen dispersión distinta y que una misma nota sitúa a los alumnos en muy distinta posición dentro del grupo.

Clase A

4,1	3,5	4,2	5,5	5,5
5,0	6,0	7,0	6,0	6,0
5,5	6,2	4,7	4,8	5,8
5,4	6,1	6,2	5,5	4,8
5,0	5,0	3,9	5,6	4,9
4,5	4,5	5,0	6,7	6,0
5,9	7,0	5,8	6,7	6,0
7,1	5,0	4,0	6,7	5,5
4,2	4,8	4,5	6,9	6,8

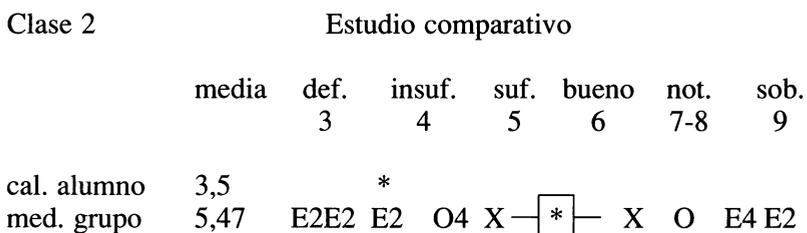
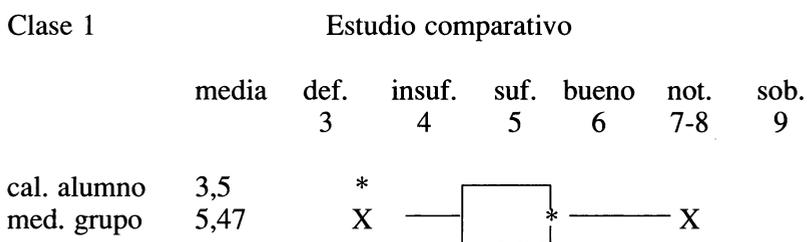
Clase B

2,0	3,4	2,7	1,8	2,7
3,2	1,5	2,4	5,0	3,6
3,2	3,7	5,5	5,0	5,5
5,0	6,0	5,9	5,8	5,8
5,9	6,0	6,8	5,0	5,0
5,0	5,0	5,0	6,3	5,5
5,0	6,2	5,8	6,8	5,9
6,8	5,9	5,9	8,9	7,9
9,9	8,7	9,0	8,5	9,9

En efecto, el alumno que tiene puntuación 3,5 puede ser el último de la clase 1 o bien ocupar un lugar próximo al primer cuartil en la clase 2, lo que se evidencia en la gráfica del estudio comparativo. Lo mismo pasaría si comparamos dos notas altas en las dos clases. Un alumno con nota 7,1 es el mejor del grupo en la clase 1 o puede estar detrás de un número relativamente grande de alumnos mejores que él en la clase 2.

Todo ello evidencia que el estudio estadístico que se viene realizando de las calificaciones escolares, que pretende ser comparativo, no sólo es insuficiente sino que también puede ser engañoso en algunos casos.

Si representamos las calificaciones con diagramas de caja se observa enseguida si el alumno está dentro del 50% central de sujetos o es un valor alejado en alguno de los extremos. Es decir, además de la información de la nota del alumno y la mediana del grupo, se da la posición del alumno dentro del grupo.



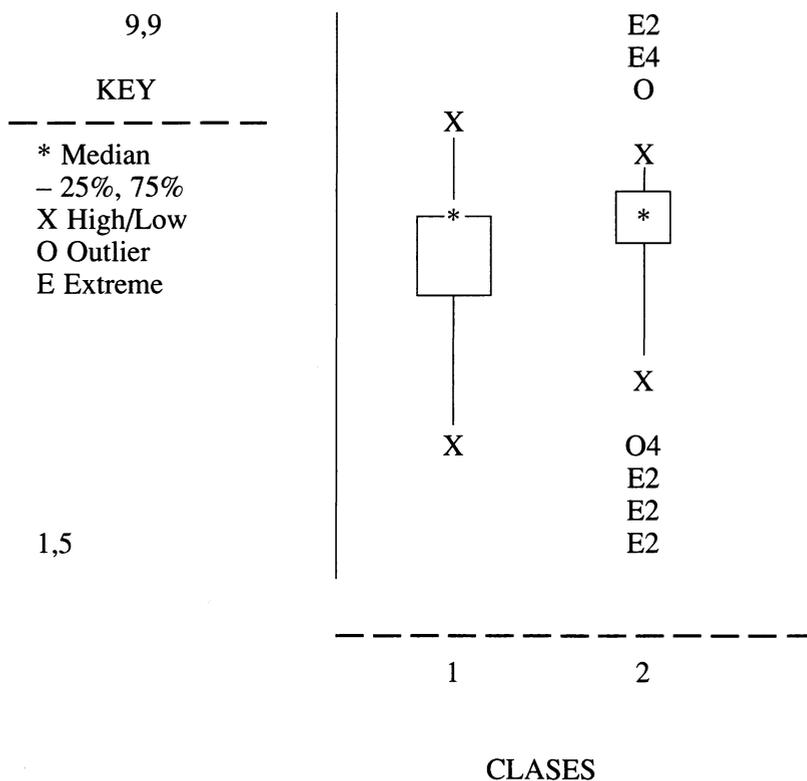
Comparación de distribuciones

Las representaciones gráficas del análisis exploratorio de datos pueden utilizarse para comparar distribuciones, en este caso las de las dos clases.

Diagrama de tallo y hojas

Clase 1		Clase 2	
3	59	1	58
4	0122	2	0477
4	55578889	3	22467
5	000004	4	
5	555556889	5	00000000055588899999
6	00000122	6	0023888
6	77789	7	9
7	001	8	579
		9	099
N = 45		N = 45	

El diagrama de tallo y hojas no esconde información. Se observa, por ejemplo, que no hay ningún cuatro en la clase 2, etc.



Ajuste de Medianas

El Análisis Exploratorio de Datos ofrece una serie de técnicas resistentes y robustas para examinar relaciones entre dos o más variables (independientes desde una perspectiva experimental) cualitativas (no necesariamente) y una variable respuesta (dependiente desde una perspectiva experimental) cuantitativa. Esta estructura de datos conocida como tablas de dos factores o diseño factorial se presenta mucho para estudiar como cada uno de los factores varía regularmente y separadamente del otro y para observar los valores que va tomando la variable respuesta según las diferentes combinaciones de los niveles y de los factores.

Estas tablas son analizadas tradicionalmente en Estadística clásica con el análisis de la varianza de dos factores.

El ajuste simple de medianas (Median Polish) descompone los efectos de la variable dependiente:

$$Y = \text{efecto común} + \text{efecto fila} + \text{efecto columna} + \text{residual}$$

Un ajuste Y para tablas de dos factores describe los datos a través de la ecuación

$$Y = X \beta + E$$

Aunque en principio el ajuste de medianas usa un modelo aditivo similar al del Análisis de la Varianza (ANOVA), ajustando éste a partir de las medianas, a través de un proceso iterativo, pone mucho énfasis en mirar y analizar los residuos.

Resumiendo, podemos decir que la técnica que introducimos ofrece, (Freixa, Salafranca, Guardia, Ferrer, Turbany, 1992) para explorar tablas de dos factores las siguientes ventajas:

- No es preciso asumir los rígidos supuestos de un modelo lineal.
- Puede analizarse con todo tipo de datos (puntuaciones directas, porcentajes, proporciones, medias, medianas, etc...).
- Puede realizarse el análisis con datos incompletos (casillas vacías).
- Es resistente.
- Explora la estructura aditiva entre las variables y mediante otras técnicas se busca la transformación más adecuada para conseguirla.
- Detecta patrones de comportamiento de los datos analizando los residuales. Mediante la descomposición de los datos intenta detectar sus patrones de comportamiento, complementando la búsqueda de estos patrones con el estudio de residuales.
- Es en general más flexible y por tanto tiene gran diversidad y riqueza de análisis y aplicaciones.

Debe advertirse, sin embargo, que el análisis de Medianas puede dar resultados un poco diferentes si el proceso de análisis se empieza por filas o por columnas.

También puede dar resultados un poco distintos según el número de iteraciones que se hagan,

pero todo ello no afectaría a las conclusiones globales extraídas.

Aunque el análisis de medianas puede usarse como técnica alternativa al ANOVA, puede plantearse como estrategia exploratoria, aportando una visión distinta y previa al análisis confirmatorio.

Se podría decir que las informaciones provenientes del análisis exploratorio resistente ayudan a señalar indicios de estructura aditiva o de aproximación a patrones o modelos, siendo ello especialmente interesante en investigaciones en Ciencias Humanas, Sociales y de la Salud.

Ejemplo:

	Clase 1	Clase 2	Clase 3	Clase 4
Asignatura A	5,8	7,2	7,3	9,1
Asignatura B	5,4	7	7,8	8,8
Asignatura C	6,2	8,8	7,4	8,1
Asignatura D	6,7	9,2	9,3	7,2

El ajuste de medianas requiere a menudo varias iteraciones. En la siguiente tabla se presentan los residuos después de dos iteraciones realizadas mediante el paquete estadístico Statgrafics.

	Clase 1	Clase 2	Clase 3	Clase 4	Efecto Asignatura
Asignatura A	0,1875	0,3625	-0,2250	0,8125	-0,275
Asignatura B	-0,2262	-0,6125	0,2250	0,4625	-0,225
Asignatura C	0,3125	0,9662	-0,4000	-0,4625	-0,000
Asignatura D	-0,1875	0,3625	0,5000	-2,3625	1
Efecto Clase	-1,9370	0,0125	-0,025	0,7375	7,82

Reproducimos el valor original

$$5,8 = 7,82 + (-0,275) + (-1,937) + 0,1875$$

En la clase 4 y en la asignatura D hay un residual alto lo cual evidencia signos de interacción. La nota promedio de la Asignatura D en la clase 4 presenta un valor menor que la mediana común, cuando lo esperado por efectos principales es que

ésta se incrementara por el hecho de ser la asignatura D y por el hecho de ser la clase 4.

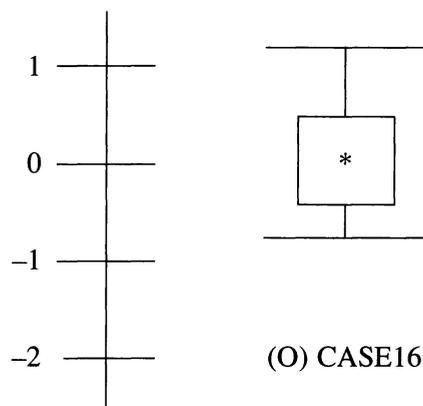


Diagrama de Caja de los residuales

CONCLUSIÓN

Con las técnicas y gráficas E.D.A. obtenemos más información de una manera rápida y sencilla. En este sencillo ejemplo los padres y tutores podrán ver la situación relativa del alumno dentro del grupo de una forma fiable. Esta información tiene gran importancia no sólo desde el punto de vista académico sino también personal.

Al Director del Centro y a los tutores estos fáciles cálculos le servirán para comparar los grupos de cada curso, que en principio se suponen homogéneos y sin grandes diferencias. Informaciones tales como que hay un tutor que no pone cuatros o notas altas, o que los alumnos menos capacitados o los mejores están en un mismo grupo, etc. se desprenden con sólo mirar los gráficos E.D.A.

Finalmente y desde un punto de vista más crítico y amplio, este ejemplo nos lleva a reflexionar sobre la utilidad y representatividad de los estadísticos clásicos que en algunos casos no son los más adecuados. Sirva el estudio, para que el usuario estadístico tome conciencia de que hay otras muchas pruebas estadísticas adecuadas a las características de cada estudio, complementarias o alternativas a las clásicas.

Tukey (1977) afirma en su libro que el E.D.A. es un *trabajo de detective numérico* para evitar confundir, mentir o cometer errores al utilizar la Estadística.

Las técnicas presentadas en este artículo son algunas de las que el E.D.A. ofrece siendo todas ellas muy útiles para analizar datos en varios contextos.

BIBLIOGRAFÍA

- BATISTA, J. M. & VALLS, M. (1985a): Nuevas técnicas de análisis estadístico de datos: Tabulación y síntesis numérica *Análisis Exploratorio de Datos*. *Qüestió*, 9, 2, 105-119.
- BATISTA, J. M. & VALLS, M. (1985b): Técnicas gráficas en *Análisis Exploratorio de Datos*. *Qüestió*, 9, 3, 163-176.
- ERICKSON, B. & NOSANCHUK, T. (1979): *Understanding Data*. Open Univ. Press: Milton Keynes.
- FREIXA, M.; SALAFRANCA, LI.; GUARDIA, J.; FERRER, R. & TURBANY, J. (1992): *Análisis Exploratorio de Datos: Nuevas Técnicas Estadísticas*. Barcelona: PPU.
- HOAGLIN, D.; MOSTELLER, F. y TUKEY, J. W. (1983) (Eds.): *Understanding Robust and Exploratory Data Analysis*. New York: John Wiley & Sons.
- HORBER, E. (1991): *Manual del paquete estadístico EDA*. Faculté des Sciences Politiques. Ginebra.
- HARTWIG, F. y DEARING, B. R. (1979): *Exploratory Data Analysis*. London: Sage.
- TUKEY, J. W. (1977): *Exploratory Data Analysis*. Reading, Massachussets: Addison-Wesley.
- VELLEMAN, P. F. y HOAGLIN, D. C. (1981): *Applications, Basics and Computing of Exploratory Data Analysis*. Boston: Duxbury.

PROBLEMAS FUNDAMENTALES DEL ANÁLISIS LOGARÍTMICO LINEAL (Y II): CASILLAS VACÍAS Y CASILLAS EXTREMAS

por

Ana Delia Correa Piñero

Dpto. Didáctica e Investigación Educativa. Universidad de La Laguna

RESUMEN

En este trabajo abordamos dos problemas que pueden presentarse en el análisis de tablas de contingencia multidimensionales: el tema de las casillas vacías —en su doble faceta de ceros muestrales y ceros estructurales— y el de las casillas y categorías extremas. Para ambos casos, se analizan sus repercusiones en el proceso y resultados del ajuste de un modelo logarítmico lineal y se ofrecen diversas soluciones, analizando sus ventajas e inconvenientes.

Descriptores: Análisis de modelos logarítmico lineales; casillas vacías; casillas y categorías «extremas»; análisis de residuales en tablas de contingencia.

ABSTRACT

Two problems that can appear in multi-way contingency tables analysis are treated in this paper. First one, empty cells in its double facet: random zeros and structural zeros. Second one, extreme (outliers) cells and strata. Both cases their effects into processes and outcomes of fitting a log-linear model are analyzed. Likewise different solutions are offered, analyzing their advantages and disadvantages.

Key words: Log-Linear Models Analysis; empty cells; extreme (outliers) cells and strata; analysis of residuals in cross-classified tables.

I. INTRODUCCIÓN

Con este artículo terminamos la serie de «problemas fundamentales» que puede encontrar el investigador al analizar un conjunto de datos mediante el ajuste de modelos logarítmico lineales. Así, en CORREA (1991a) tratamos diversos problemas relacionados con el «colapsamiento» de variables (recombinar categorías, categorizar datos continuos, etc.). Otros problemas de interés, relacionados con el tamaño de la muestra, pueden encontrarse en CORREA (1991b). En esta ocasión nos centraremos en los problemas que supone para la fase de análisis la presencia de casillas vacías y casillas extremas. Este último tema también es conocido en la literatura como *análisis de residuales*.

Se presupone que el lector está familiarizado con el análisis logarítmico lineal en lo que respecta a la elaboración de tablas de contingencia multidimensionales, selección de modelos, ajuste de modelos y cálculo de parámetros y errores típicos. Pueden consultar una introducción al tema y otras referencias de interés en CORREA (1991c).

II. EL PROBLEMA DE LAS CASILLAS VACÍAS: CEROS MUESTRALES Y CEROS ESTRUCTURALES

Se dice que una casilla está vacía cuando tiene frecuencia cero. No hay ningún sujeto en una determinada combinación de categorías de las variables implicadas. Esto supone problemas, ya que la base del análisis logarítmico lineal está en el cálculo de razones (*ratios*) y la presencia de un cero en el denominador produciría una indeterminación. Este hecho puede producirse por dos motivos, que dan lugar a dos tipos de casillas vacías: los *ceros muestrales* y los *ceros estructurales*.

Cuando en la muestra no se obtiene ningún caso para determinadas combinaciones de categorías se producen casillas con frecuencia cero, pero sin que eso signifique que no existan esas combinaciones en la población. Simplemente, en la muestra obtenida no se presenta ningún caso. Cuanto más pequeña sea la muestra, aumenta la probabilidad de que alguna(s) casilla(s) quede(n) sin cubrir. Por el contrario, al crecer el tamaño de la muestra aumenta la probabilidad de que todas las casillas resulten ocupadas. A este tipo de casillas se les llaman *ceros muestrales o aleatorios* (GOLDSTEIN y DILLON, 1980).

Se han apuntado varias soluciones para este problema. GOODMAN (1970) propone añadir un pequeño valor constante (normalmente 0,5) a cada casilla antes de efectuar los cálculos. Éste es un procedimiento conservador que puede subestimar los parámetros y su significación: la constante añadida aumenta las frecuencias esperadas, lo cual reduce el valor de los estadísticos χ^2 y G^2 que son los que se utilizan para comprobar la calidad de ajuste del modelo (BROWN, 1983).

Ya que en los cálculos de las razones, como dijimos, la división entre 0 daría «indeterminado», otra solución sería definir, de forma arbitraria, que $0/0=0$ (FIEN-

BERG, 1977). En ese caso debe tenerse precaución con dos aspectos: uno, la posible aparición de frecuencias esperadas muy pequeñas que ponen en entredicho la aplicabilidad de los estadístico de ajuste; el otro, si los cálculos se hacen mediante algún programa computarizado, asegurarse que esa solución es factible mediante las rutinas programadas o se corre el riesgo de que se «aborte» el análisis, como sucede, por ejemplo, al intentar invertir una matriz cuyo determinante es cero.

Otra posible solución, la más obvia pero no siempre posible por motivos diversos, es aumentar el tamaño de la muestra lo suficiente como para que se eliminen los ceros (KNOKE y BURKE, 1982). Se entiende que se vuelve a retroceder a la fase de obtención de datos hasta que desaparezca el problema.

Finalmente, queda una cuarta alternativa: colapsar o recombinar categorías, de manera que la casilla vacía quede «integrada» en otra casilla con frecuencias suficientes. Las posibles repercusiones del colapsamiento en el ajuste del modelo y en el valor de los parámetros se analizaron en el artículo anterior (CORREA, 1991a).

El segundo tipo de casilla vacía se produce cuando en la tabla hay ciertas categorías que no tienen referente empírico y, por tanto, tendrán frecuencia cero aunque se trabaje con toda la población. Son los *ceros lógicos*, llamados también *ceros estructurales*, ya que son consecuencia de la estructura del problema (FAGEN y MANKOVICH, 1980). Esta clase de ceros puede surgir por:

a) *Diseño del muestreo*. Es decir, que el investigador en su diseño del muestreo decida suprimir cierta casilla por razones teóricas. En realidad, este caso sí tendría referente empírico, pero sería teóricamente irrelevante.

b) *Secuencia ordinal de las variables*. Por ejemplo, en una tabla «Edad x Status Familiar», es poco probable encontrar la categoría «abuelos» en los primeros órdenes (intervalos) de la variable edad.

c) *Inconsistencia definicional*. Es el caso más neto de cero lógico. Por ejemplo, en una tabla «Sexo x Tipo de operación quirúrgica» es imposible encontrar referentes empíricos correspondientes a la casilla «mujer-próstata».

Por cualquiera de estas razones, pues, surgen las denominadas tablas de contingencia incompletas o truncadas. GOODMAN (1968) se ha interesado por este tipo de tablas, para el caso de dos y tres variables, formulando una técnica para su análisis conocida como prueba de cuasi-independencia. En una tabla con casillas vacías, la cuasi-independencia es una forma de «no asociación» entre las variables, pero referida sólo a la porción de la tabla donde no haya ceros. Es decir, es una forma de *independencia condicional* donde se modifican los grados de libertad restando el número de ceros estructurales existentes en la tabla. En una tabla $I \times J$, con Z casillas 0, los grados de libertad son: $(I-1)(J-1)-Z$. Así, a partir de las casillas disponibles, se obtienen estimaciones de máxima verosimilitud para las casillas vacías. Éstas son reemplazadas por dichos valores estimados. Luego, se calculan los valores esperados de todas las casillas como si la tabla fuese entera (véase BROWN, 1975). FIENBERG (1972) generaliza las ideas básicas a tablas mayores, formulando lo que él denomina modelos cuasi logarítmico lineales.

Por otra parte, SMITH (1973) señala la utilidad de las tablas incompletas para

abordar cierto tipo de problemas, por ejemplo: permiten prescindir de respuestas incorrectas cuando se desea analizar sólo las correctas, o de los «acuerdos» entre jueces cuando sólo interesa analizar los «desacuerdos» y similares.

Personalmente, en la práctica hemos encontrado un «tercer tipo» de casilla vacía, a medio camino entre el cero muestral y el estructural. En una encuesta a profesores de EGB sobre medios didácticos (CORREA y AREA, en prensa), utilizamos un cuestionario que incluía, entre otros, un ítem donde se pedía que indicaran la importancia que le concedían a una serie de elementos del libro de texto. Los elementos textuales en cuestión eran los siguientes:

- 1) Su relación con los Programas Oficiales.
- 2) El tipo de actividades que propone.
- 3) El lenguaje utilizado.
- 4) Los aspectos formales (colorido, tamaño, ilustraciones, etc.).
- 5) Su adecuación al nivel de conocimientos de los alumnos.
- 6) Lo innovador de su planteamiento.
- 7) La extensión con que trata los contenidos.
- 8) Su relación con el contexto.
- 9) El método de enseñanza que se desprende el texto o de su guía didáctica.

La importancia se reflejaba en una escala ordinal de 4 niveles (desde «ninguna» a «mucha»). Cada elemento textual, pues, representa una variable distinta, con categorías ordenadas. No resultó difícil encontrar sujetos que, para un elemento aislado, se repartiesen a lo largo de todo el continuo de valoración. El cruce de varios elementos (variables) produce tanto casillas con varias valoraciones simultáneas de «mucha importancia» como casillas donde se reflejan varias valoraciones simultáneas de «ninguna importancia». Pues bien, estas últimas casillas resultaron, en su mayoría, vacías. No se trataba de ceros estructurales, en sentido estricto. Siempre cabe pensar en la existencia de profesores/as en la población que puedan ubicarse en esas casillas, de forma que al aumentar el tamaño de la muestra encontrásemos algún «referente empírico» apropiado. Pero, debido a la formulación misma del contenido del ítem, pensamos que serían casos muy raros, de tal forma que ese necesario aumento de muestra debería ser muy considerable, lo cual hace que sea también difícil calificarlas de ceros muestrales estrictos. Los elementos textuales citados, innegablemente, son apreciativos. En cualquier caso, neutros; pero no peyorativos. De tal forma que la negación simultánea de la importancia de varios de ellos conjuntamente es un «caso raro». Y se va haciendo progresivamente más raro a medida que la tabla de contingencia incluya más elementos textuales. Es decir, es frecuente encontrar profesores/as que dan poca importancia a un elemento, considerado aisladamente (análisis univariado). Un poco menos frecuente, pero no raro, encontrar esa misma valoración referida a un par de elementos a la vez (tabla bivariada). Bastante raro, si la valoración hace referencia a tres elementos conjuntamente (tabla trivariada). Etc, etc.

Podría aducirse que este tipo de ceros sería mejor considerarlos como ceros

estructurales del tipo a), es decir, fruto del diseño del muestreo. Pero, en primer lugar, no se decidió a priori suprimir esas casillas. Y, en segundo lugar, en este caso no era teóricamente irrelevante, sino todo lo contrario, la constatación de su existencia.

En suma, en una aplicación concreta de esta técnica de análisis, podemos encontrar que «algo» en nuestras tablas de contingencia no se ajusta a los modelos standard. En ese caso, debería tomarse la decisión que se considere más apropiada, de acuerdo con las peculiaridades del estudio. Nosotros, en este caso y por las razones expuestas, decidimos recurrir a dos soluciones y posteriormente comparar los resultados. Una, tratar dichas casillas como ceros muestrales y seguir el criterio sugerido por Goodman de añadir una constante, en lugar de aumentar la muestra a la búsqueda, probablemente sin éxito, de referentes empíricos reticentes. La segunda solución fue combinar categorías adyacentes (véase artículo anterior sobre colapsamiento en esta misma revista) ya que las categorías de las variables eran ordenadas y su reagrupación, aunque disminuía el recorrido de la variable, no alteraba la ordenación subyacente. En la mayoría de los casos (tablas) fue preciso combinar ambos recursos. Los resultados que se ofrecen en la referencia citada (CORREA y AREA, en prensa) son precisamente estos últimos, ya que así obteníamos el ajuste con modelos más parsimoniosos.

III. EL PROBLEMA DE LAS CASILLAS EXTREMAS

En ocasiones, puede suceder que una frecuencia de casilla, incluso de una categoría entera, se aleje excesivamente de lo que cabría esperar a partir del modelo de referencia. Dicho de otra forma, la frecuencia esperada es excesivamente grande o excesivamente pequeña. Esto puede llevar a una impresión distorsionada de las relaciones entre variables, ya que «las casillas extremas pueden sesgar los valores esperados lo suficiente como para impedir que se hagan inferencias sobre el resto de casillas» (FAGEN y MANKOVICH, 1980:1.020). Es decir, la presencia de un valor extremo no sólo afecta la casilla donde se produce, sino a todas las casillas de la fila y columna a la que pertenece y, en último término, afecta la significación del modelo.

Es conveniente, pues, disponer de algún procedimiento para detectar casillas y/o niveles extremos, de la misma forma que otros procedimientos estadísticos disponen de sistemas para detectar «sujetos extremos», como la D^2 de Mahalanobis, el método de Comrey, el análisis de los residuos en la regresión, etc. (véase GAVIRIA, 1988). Esto podría explicar por qué en ocasiones se producen malos ajustes en modelos que, según todos los indicios eran muy prometedores. Estos indicios pueden venir dados, por ejemplo, por las pruebas previas de asociación parcial y marginal y su determinación de efectos «innecesarios», «necesarios» e «incierto».

La base del análisis logarítmico lineal radica en la comparación de frecuencias observadas y esperadas. Una «anomalía» se detectaría entonces como una discrepancia, o residual, proporcionalmente grande entre ambos tipos de frecuencias. No

debe confundirse «casilla extrema» con «casilla de mayor frecuencia». La «extremosidad» viene dada por un mayor o menor alejamiento entre una frecuencia esperada y su correspondiente observada. Puede ser que, para parecerse a lo que se espera *bajo cierto modelo*, una frecuencia absoluta tenga que ser muy pequeña o puede que tenga que ser muy grande. Además, hemos dicho «proporcionalmente grande», es decir, que depende de los datos: no es lo mismo una diferencia de 100 cuando se comparan 12.000 y 11.900 que cuando se comparan 12 y 112. En este último caso, la diferencia es vital para el ajuste o desajuste del modelo.

Para detectar estos residuales se han propuesto diversas fórmulas, que dan lugar a otros tantos tipos de residuales (HABERMAN, 1973; UPTON, 1978). Como se trata de desviaciones a partir de las frecuencias esperadas, se les suele llamar así, *desviaciones*. Existen 3 tipos de desviaciones: de Freeman-Tukey, standarizadas y standarizadas ajustadas. Las dos primeras pueden ser aplicadas a tablas bi y multi-variadas. La última, sólo a tablas de dos variables.

La desviación de Freeman-Tukey en cualquier casilla ij se define como:

$$d_{ij}(F-T) = \sqrt{(o_{ij})} + \sqrt{(o_{ij}+1)} - \sqrt{(4e_{ij}+1)} \quad (\text{Ec.1})$$

siendo o las frecuencias observadas en la casilla ij y e las frecuencias esperadas.

La interpretación de los valores obtenidos se basa en la similitud de estas desviaciones con las puntuaciones típicas. A partir de la distribución normal unitaria se determina el porcentaje correspondiente a esa puntuación (desviación), valorando así la magnitud del residual. No obstante, para poder hacer esta analogía con puntuaciones z , es preciso que los datos se hayan obtenido mediante una técnica de muestreo específica: la distribución Poisson.

Para una casilla ij cualquiera se define la desviación estandarizada o típica como:

$$d_{ij}(\text{est}) = (o_{ij} - e_{ij})/\sqrt{e_{ij}} \quad (\text{Ec.2})$$

Son los mismos elementos que intervienen en la conocida fórmula de χ^2 , con la excepción de la raíz cuadrada. Para su análisis se recurre igualmente a la distribución normal.

Finalmente, las desviaciones típicas ajustadas son desviaciones estandarizadas, ajustadas de tal forma que tengan varianza 1 cuando los datos provengan de una distribución multinomial. El inconveniente, como decíamos, es que sólo se aplica a tablas bivariadas.

Su fórmula, dada la casilla ij , es:

$$d_{ij}(\text{ajs}) = \frac{o_{ij} - e_{ij}}{\sqrt{[e_{ij}(1-f_{i+}/N)(1-f_{+j}/N)]}} \quad (\text{Ec.3})$$

Los nuevos elementos, f_{i+} y f_{+j} , representan el total marginal de fila y columna, respectivamente. También se interpretan sus valores a partir de la tabla de la distribución normal.

Adicionalmente, habría otra medida de desviación: la diferencia entre las frecuencias observadas y las esperadas (o-e); este índice, sin embargo, es muy simple y no tiene base para una interpretación clara, como los anteriores.

Interpretación de los residuales

Al comienzo de este apartado comentábamos la importancia del análisis de casillas extremas para ver por qué un modelo hipotetizado no había ajustado, como se esperaba. Vamos a detallarlo a través de un ejemplo con los datos ficticios de una tabla tridimensional: SEXO x PERÍODO DE TIEMPO x CARRERA (véase Tabla 1). Todos los análisis se han efectuado mediante el programa 4F de BMDP.

TABLA 1
CARRERA x SEXO x PERÍODO DE TIEMPO (frecuencias observadas)

CARRERA	PERÍODO	SEXO	
		Hombre	Mujer
Ciencias	1960-70	69	12
	1971-80	124	34
Humanas	1960-70	85	113
	1971-80	68	116
Técnicas	1960-70	105	36
	1971-80	136	98

Inicialmente, aplicamos las pruebas de asociación parcial y marginal a los datos de la Tabla 1. Los resultados indican que las interacciones SC (SEXO y PERÍODO DE TIEMPO) y PC (PERÍODO DE TIEMPO y CARRERA) son efectos «definitivamente necesarios». La interacción SP, por su parte, resulta «incierto», ya que sólo es significativa en una de las pruebas. Sometemos a ajuste, por tanto, dos modelos: el SC,PC (que incluye sólo efectos «definitivamente necesarios») y el SC,PC,SP (que contiene los anteriores y, además, el efecto «incierto»). Los resultados del ajuste se muestran en la Tabla 2.

TABLA 2
AJUSTE DE MODELOS A LOS DATOS DE LA TABLA 1

MODELO	GL	G ²	PROB	JF ²	PROB
SC,PC	3	13,52	0,0036	13,21	0,0042
SC,PC,SP	2	2,49	0,2885	2,47	0,2907

Como se ve, el modelo SC,PC no presenta un buen ajuste: sus índices residuales (tanto ji^2 como G^2) son significativos; luego, el ajuste no lo es. En cambio, sí ajusta bien el modelo que incluye el efecto «incierto»: los residuales de SC,PC,SP no son significativos, por tanto, sí es significativo su ajuste. Sin embargo, el valor de G^2 componente correspondiente al efecto «incierto» (SP) es muy pequeño, lo que quiere decir que este efecto contribuye muy poco al ajuste del modelo SC,PC,SP. En concreto, su contribución o componente se determina mediante la sustracción de residuales de dos modelos que sólo se diferencien en el efecto de interés; sería: $13,52 - 2,49 = 11,03$. La contribución de los otros dos efectos SC y PC, calculada de la misma forma y mediante los residuales de los modelos apropiados (que no se muestran en la Tabla 2), obtuvieron valores superiores: 120,2 y 33,3, respectivamente.

¿Por qué no ajustó bien el modelo SC,PC cuando todos los indicios hacían pensar lo contrario? Analicemos los residuales de ese modelo. Ya que se trata de una tabla de 3 variables no podemos usar las desviaciones estandarizadas ajustadas; las de Freeman-Tukey tampoco serían apropiadas, ya que los datos se corresponden con un esquema de muestreo multinomial (ningún marginal está fijado de antemano) y no un esquema Poisson; las diferencias o-e, dijimos que eran poco informativas. Así, pues, analizaremos mediante las desviaciones estandarizadas si se detecta la presencia de alguna casilla extrema, que haya contribuido al desajuste de un modelo tan prometedor. Las desviaciones estandarizadas para esos datos bajo el modelo SC,PC se muestran en la Tabla 3.

TABLA 3
DESVIACIONES ESTANDARIZADAS BAJO EL MODELO SC,PC

CARRERA	PERÍODO	SEXO	
		Hombre	Mujer
Ciencias	1960-70	0,4	-0,9
	1971-80	-0,3	0,7
Humanas	1960-70	0,6	-0,5
	1971-80	-0,7	0,5
Técnicas	1960-70	1,5	-2,0
	1971-80	-1,2	1,6

Numerando las casillas de arriba a abajo y de izquierda a derecha, las que muestran unas mayores desviaciones son la 9, 10, 11 y 12. Todas se ubican en el nivel «técnicas». Sobre todo la casilla 10 muestra una desviación considerable (-2,0). Una puntuación típica de -2,0 es un valor muy extremo. Se corresponde, en

la distribución normal, con un área de 0,0228. Es decir, que sólo un 2,28% de observaciones elegidas a partir de la distribución normal unitaria estaría *por azar* a esa distancia o mayor de la media 0. La probabilidad, pues, de que esta distancia se deba al azar es muy pequeña. Es una desviación significativa, que contribuye de forma más acusada que las otras al desajuste que mostró el modelo SC,PC.

El mismo cálculo para el modelo SC,PC,SP se muestra en la Tabla 4.

TABLA 4
DESVIACIONES ESTANDARIZADAS BAJO EL MODELO SC,PC,SP

CARRERA	PERÍODO	SEXO	
		Hombre	Mujer
Ciencias	1960-70	-0,0	0,0
	1971-80	0,0	-0,0
Humanas	1960-70	-0,5	0,5
	1971-80	0,6	-0,5
Técnicas	1960-70	0,5	-0,8
	1971-80	-0,4	0,5

En este caso las diferencias se reducen, sobre todo en las casillas 9, 11 y 12. En la 10 también se reduce (esto es lógico, ya que este modelo sí ajustó), pero, con todo, esta casilla sigue mostrando la mayor desviación de la tabla. Es decir, es la que menos contribuye al ajuste del modelo.

¿Qué conclusión sacar a partir de esto? Simplemente, que el nivel «técnicas» combinado con el resto de niveles del resto de variables, muestra un comportamiento muy distinto a los demás. Es un «caso muy raro», difícilmente explicable a partir del modelo SC,PC, aunque «menos raro» a la luz del modelo SC,PC,SP.

Esto también podría observarse con el cálculo de parámetros: los parámetros de menos valor (y menos significativos) para el modelo SC,PC,SP se obtendrán en aquellos términos donde esté implicado el nivel «técnicas». En concreto, una vez calculados, obtenemos que para el nivel «técnicas» al parámetro de interacción SEXO-CARRERA le corresponde un valor (λ) de 0,033; y al de la interacción CARRERA-PERÍODO DE TIEMPO un valor de 0,077. Estas cifras suponen los menores valores de todo el conjunto de parámetros y, como se sabe, cuanto más cercanos a cero están los parámetros, menor es su importancia en el modelo de referencia.

A lo largo de este ejemplo, vemos como los resultados de las distintas fases del análisis van armonizando, complementándose, para dar una idea global del esquema que subyace a los datos: las pruebas de asociación parcial y marginal nos aconsejan

dos determinados modelos para su ajuste; probamos ambos modelos y uno de ellos ajusta y el otro no; sin embargo, el valor G^2 componente del término que diferencia a un modelo de otro es muy reducido, contribuye poco al ajuste; si esto es así, ¿por qué no ajustó también el otro modelo?; buscamos si existen casillas extremas que hayan podido impedirlo y, efectivamente, las encontramos; y coinciden, precisamente, con los parámetros de menor importancia en el modelo. El buen concierto de unos y otros resultados llega a producir un placer casi estético.

¿Qué decisión tomamos con estas casillas extremas y qué influencia puede tener en el análisis la decisión tomada? Ya que «técnicas» contribuye muy poco, parece lógico pensar que es muy probable que si se reduce la variable CARRERA a dos niveles, eliminando «técnicas», se mantenga el esquema de relaciones entre SEXO y CARRERA y PERÍODO DE TIEMPO y CARRERA. Vamos a comprobarlo.

Solución extrema para los datos extremos: eliminación de categorías

Supongamos que a la hora de obtener datos sólo hemos encuestado a hombres y mujeres, de los dos períodos, que estudian ciencias o humanas. Ningún técnico. La fila «técnicas» desaparecería de la Tabla 1 dando lugar a la Tabla 5.

TABLA 5
TABLA 1 SIN LA CATEGORÍA «TÉCNICAS» (FRECUENCIAS OBSERVADAS)

CARRERA	PERÍODO	SEXO	
		Hombre	Mujer
Ciencias	1960-70	69	12
	1971-80	124	34
Humanas	1960-70	85	113
	1971-80	68	116

La diferencia entre el N anterior (996) y el actual (621) no distorsiona lo que queremos averiguar. Al contrario, lo refuerza, ya que cuanto mayor sea N, aumenta la probabilidad de encontrar ajustes significativos, hasta con modelos complejos, como el presaturado, y viceversa. El ajuste del modelo SC,PC,SP para los nuevos datos es:

TABLA 6
AJUSTE DEL MODELO SC,PC,SP PARA LOS DATOS DE LA TABLA 5

MODELO	GL	G^2	PROB	ji^2	PROB
SC,PC,SP	1	0,24	0,6246	0,24	0,6265

No sólo se mantiene el ajuste que ya mostró la tabla original, sino que se mejora: los índices residuales (G^2 y ji^2) pasan de 2,49 y 2,47, respectivamente, a ser de 0,24 en ambas pruebas. Las desviaciones estandarizadas bajo este modelo se muestran en la Tabla 7.

TABLA 7
DESVIACIONES ESTANDARIZADAS DE LA TABLA 5 BAJO EL MODELO SC,PC,SP

CARRERA	PERÍODO	SEXO	
		Hombre	Mujer
Ciencias	1960-70	0,1	-0,3
	1971-80	-0,1	0,2
Humanas	1960-70	-0,1	0,1
	1971-80	0,1	-0,1

Vemos que las desviaciones disminuyen considerablemente con respecto a las obtenidas para los datos originales. Eliminando «técnicas», el modelo que ya había ajustado en la tabla original mejora su ajuste. Veamos en la Tabla 8 qué sucede ahora con el modelo tan prometedor, que no llegó a ajustarse.

TABLA 8
AJUSTE DEL MODELO SC,PC A LA TABLA 5

MODELO	GL	G^2	PROB.	ji^2	PROB.
SC,PC	2	3,02	0,2208	2,97	0,2270

En esta ocasión, el modelo sí ajusta. Y sus desviaciones estandarizadas muestran la desaparición de casillas extremas, tal como se muestra en la Tabla 9.

TABLA 9
DESVIACIONES ESTANDARIZADAS DE LA TABLA 5 BAJO EL MODELO SC,PC

CARRERA	PERÍODO	SEXO	
		Hombre	Mujer
Ciencias	1960-70	0,4	-0,9
	1971-80	-0,3	0,7
Humanas	1960-70	0,6	-0,5
	1971-80	-0,7	0,5

A pesar de estos resultados, quizá parezca excesivamente riguroso suprimir de un plumazo una categoría entera. Al fin y al cabo, no todas las casillas donde intervenía el nivel «técnicas» eran igualmente extremas. Para el modelo SC,CP las casillas más extremas (9, 10, 11, 12) tenían desviaciones de 1,5, -2,0, -1,2 y 1,6, respectivamente. Quizá bastaría con eliminar la casilla 10.

La solución conciliadora: eliminación de casillas

Para aplicar esta solución menos drástica, es preciso dejar la tabla tal y como estaba, excepto que se anula una casilla, como si su frecuencia fuera cero. Esto supone obtener una tabla incompleta, en la cual se ajusta el modelo mediante una prueba de cuasi-independencia, como explicamos en el apartado de casillas vacías.

El programa 4F de BMDP puede, a partir de un modelo que no ha ajustado bien, ir identificando paso a paso las casillas más extremas (en cada paso, una casilla) y recalculando el ajuste eliminándolas por turno. Cesa el proceso cuando consigue un buen ajuste. El sistema, por tanto, no funciona con un modelo que ajuste bien desde el principio. Veamos en la Tabla 10 qué sucede al aplicar este procedimiento a los datos originales.

TABLA 10
AJUSTE DEL MODELO SC,PC CON ELIMINACIÓN AUTOMÁTICA DE CASILLAS EXTREMAS (DATOS TABLA 1)

PASO	MODELO X ²	SC,PC GL	PROB	DESVIACIÓN MÁXIMA	ENCONTRADA EN LA CASILLA
0	13,52	3	0,00363		
1	3,02	2	0,22079	-2,026	TÉCNICAS 1960-70 MUJER

En el paso 0 se ajusta el modelo con la tabla completa y se identifica la casilla con una desviación estandarizada máxima. Como ya sabíamos, es la casilla 10. Luego, paso 1, elimina esa casilla y calcula el ajuste de la tabla incompleta. Ahora los datos observados sí se ajustan al modelo y el proceso se detiene.

Efectivamente, suprimir el nivel «técnicas» entero, fue una decisión bastante drástica. Por ello, se aconseja utilizar primero este proceso paso a paso con casillas extremas y ver si así se consigue un buen ajuste. Si, después de todo, fuera preciso eliminar un nivel entero, se eliminaría en última instancia.

IV. CONSIDERACIONES FINALES

Hay una deducción, que quizá resulte obvia al lector, con respecto a las casillas «raras» (vacías o extremas) y los recursos más habituales para solventarlas. Tanto el colapsamiento como la eliminación de casillas o niveles extremos, favorecen el ajuste de modelos más parsimoniosos, menos complejos. En principio, esto parece deseable, pero podría alegarse que colapsando por aquí y eliminando por allá nos quedaremos con «la norma», cuando en ocasiones los casos «raros» pueden ser teóricamente interesantes, incluso deseables si se hipotetizan relaciones complejas, lo cual es muy probable en investigación educativa (en investigación social, en general).

En este punto, podemos citar la recomendación de GAVIRIA (1988) con respecto a los casos extremos: si se trata de investigación básica, con una muestra representativa, los casos extremos deben incluirse en el análisis, ya que son elementos representativos de grupos que, aunque poco numerosos, tendrían más probabilidad de aumentar a medida que se incrementase el tamaño de la muestra. Por el contrario, en investigación operativa, para tomar decisiones locales, sería preferible eliminar esos sujetos del análisis y considerarlos individualmente. No obstante, al margen de la utilidad de esta sugerencia, hay que recordar que en un análisis logarítmico lineal no siempre es aplicable la consideración de «casos (casillas) extremas» como «casos poco numerosos». Aquí una casilla es extrema no por su frecuencia absoluta, observada, sino por su discrepancia con la frecuencia esperada bajo cierto modelo. Así, (tenga una frecuencia alta o baja) puede ser extrema para ciertos modelos y no extrema para otros. Por tanto, la decisión está, sobre todo, en manos del investigador y del modelo que adopte. Con estos dos artículos sobre «problemas fundamentales» hemos querido contribuir a que las decisiones que tome al respecto en un análisis logarítmico lineal estén un poco mejor fundamentadas e informadas.

REFERENCIAS

- BROWN, M. B. (1975): The asymptotic standard errors of some estimates of uncertainty in the two-way contingency table. *Psychometrika*, 40(3), 291-296.
- BROWN, M. B. (1983): PF4. Two-way and Multi-way Frequency Tables - Measures of Association and the Log-Linear Model (Complete and Incomplete Tables). En W. J. Dixon (Ed.) *BMDP Statistical Software*. California: University of California Press.
- CORREA, A. D. (1991a): Problemas fundamentales del análisis logarítmico lineal (I): el colapsamiento de variables y su influencia en el ajuste de modelos. *Revista de Investigación Educativa*, nº 18, 2º semestre.
- CORREA, A. D. (1991b): Reglas prácticas en torno al tamaño de la muestra para el ajuste de modelos logarítmico lineales. *Curriculum*, Nº Extra 1/2, 365-368.
- CORREA, A. D. (1991c): Estudios multivariados con datos nominales: aportaciones del análisis logarítmico lineal. *Curriculum*, nº 3, 35-52.
- CORREA, A. D. y AREA, M.: ¿Qué opinan los profesores de EGB sobre el uso del libro de texto en las escuelas? *Curriculum*, nº 4, en prensa.

- FAGEN, R. M. & MANKOVICH, N. J. (1980): Two-act transitions, partitioned contingency tables, and the «significant cells» problem. *Animal Behaviour*, 28(4), 1.017-1.023.
- FIENBERG, S. E. (1972): The analysis of incomplete multiway contingency tables. *Biometrics*, 28, 177-202.
- FIENBERG, S. E. (1977): *The analysis of cross-classified categorical data*. Cambridge, Massachusetts: MIT Press.
- GOLDSTEIN, M. & DILLON, W. R. (1980): A measure of separability and random zeros in statistical classification. *Multivariate Behavioral research*, 15(1), 57-71.
- GAVIRIA, J. L. (1988): La detección de «outliers» en los análisis multivariantes. En I. Dendaluze (Coord.): *Aspectos metodológicos de la investigación educativa* (pp. 264-271). Madrid: Narcea.
- GOODMAN, L. A. (1968): The analysis of cross-classified data. *Journal of the American Statistical Association*, 63, 1.091-1.131.
- GOODMAN, L. A. (1970): The multivariate analysis of qualitative data: interactions among multiple classifications. *Journal American Statistical*, 65, 226-256.
- HABERMAN, S. J. (1973): The analysis of residual in cross-classified tables. *Biometrics*, 29, 205-220.
- KNOKE, D. & BURKE, P. J. (1982): *Log-Linear Models*. Beverly Hills, CA: Sage Pubns.
- SMITH, J. E. (1973): On tests of quasi-independence in psychological research. *Psychological Bulletin*, 80(4) 329-333.
- UPTON, G. (1978): *The analysis of cross-tabulated data*. New York: Wiley.

Autor: Catalina M. Alonso García

Dirección: Argos, 31
28220 Majadahonda (Madrid)

Director: Arturo de la Orden Hoz
Dpto.: Métodos de Investigación y Diagnóstico en Educación
Dirección: Paseo de las Moreras s/n. Ciudad Universitaria.
28040 Madrid

Centro: Facultad de Educación Universidad Complutense

Descriptorios

Estilos de Aprendizaje; Enseñanza Universitaria

Bibliografía fundamental

CANFIELD, A. A. (1988): *Canfield Learning Styles Inventory*. Humaniscs Media, Detroit, Michigan.
 DUDA, R. y RILEY, P. (1990): *Learning Styles*, Press Universitaires de Nancy, Nancy.
 DUNN, P.; DUNN, K. y PRICE, G. (1979): *Learning Style Inventory*, Price Systems, Lawrence, Kansas.
 DUNN, R.; DUNN, K. (1984): *La Enseñanza y el Estilo Individual de Aprendizaje*, Anaya, Madrid.
 HONEY, P. y MUMFORD, A. (1986): *Using our Learning Styles*, Peter Honey, Berkshire.
 JUCH, B. (1987): *Desarrollo personal*, Limusa, México.
 KEEFE, J. W. (1982): *Assessing Student Learning Styles*, ERIC, ED. 227565.
 KEEFE, J. W. (1988): *Profiling and Utilizing Learning Style*, NASSP, Reston, Virginia.
 KOLB, D. (1976): *The Learning Style Inventory: Technical Manual*, McBer and Company, Boston.
 ROSS, D. (1985): *Learning Styles*, ERIC, ED. 269201.

Problema investigación

Existen diferentes maneras de aprender que dependen tanto de los rasgos cognitivos, afectivos y fisiológicos de cada uno, como de la forma en que se ha recibido la Formación (educación e instrucción) y las experiencias de aprendizaje vividas.

El núcleo de la investigación consistía en saber si se podía averiguar las diferentes maneras de aprender de nuestros estudiantes. Quiénes han estudiado anteriormente este problema. Qué Instrumento de medición es el más adecuado. Diseñar dieciocho Cuestiones socioacadémicas con las que posiblemente puedan tener relación los Estilos de Aprendizaje, para contrastar si tienen o no diferencias significativas con los cuatro Estilos de Aprendizaje: Activo, Reflexivo, Teórico y Pragmático.

Las dieciocho Hipótesis se relacionan con rasgos académicos y sociales:

Académicos: Facultad (y Grupos de Facultades), Curso, Nota media del curso anterior y de ingreso en la Facultad, Carrera elegida en primer lugar y en segundo lugar, Área de Bachillerato donde tenía las notas más altas y las notas más bajas, Ha convalidado estudios, Además de estudiar trabaja.

Sociales: Edad, Sexo, Estudios del padre y de la madre, Profesiones del padre y de la madre, Número de hermanos, Ciudad de origen.

Muestra y método de muestreo

N=1.371 alumnos de cuarto o quinto curso de carrera de veinte y cinco Facultades/ Escuelas de las Universidades Complutense y Politécnica de Madrid.

CAT ^(a)	ANO ^(a)	CLASIFICACIÓN ^(a)
D	1 9 9 1	5 8 9 9 9

N.º CITAS ^(a)	N.º PÁGINAS
8 2 8 1 3 1 5	

De forma aleatoria se aplicó el Cuestionario HONEY-ALONSO de Estilos de Aprendizaje: CHAEA a un grupo de cada una de las veinte y cinco Facultades/Escuelas siempre por la misma persona, la autora de esta investigación.

Metodología del trabajo

1. Amplia búsqueda bibliográfica, manual e informática, sobre el tema.
2. Selección de contenidos específicos, más pertinentes.
3. Plantear las Hipótesis referentes a las dieciocho cuestiones socioacadémicas.
4. Fundamentar, explicitar, clasificar, exponer las distintas teorías/Cuestionarios seleccionados de Estilos de Aprendizaje.
5. Seleccionar un Cuestionario: LSQ de P. HONEY aplicado a profesionales inglés.
6. Traducir y adaptar LSQ al contexto español.
7. Aplicar el nuevo Cuestionario —CHAEA— a 1.371 estudiantes universitarios.
8. Probar la Fiabilidad y Validez del nuevo Cuestionario CHAEA.
9. Exponer baremos que sirvan de referencia en las aplicaciones futuras del Cuestionario.
10. Indicar algunas propuestas de optimización para los alumnos que han tenido Preferencia Baja o Muy Baja en un Estilo de Aprendizaje determinado.
11. Contraste de las Hipótesis planteadas.
12. Conclusiones y prospectiva del significado de los Estilos de Aprendizaje en la Enseñanza Universitaria.
13. Presentar una amplia bibliografía específica. 828 títulos sobre Estilos de Aprendizaje.

Técnicas de análisis

Para probar la fiabilidad del cuestionario CHAEA:

- Pruebas Alfa de Cronbach en los cuatro Estilos de Aprendizaje: Activo, Reflexivo, Teórico y Pragmático.

Como indicadores de validez del cuestionario CHAEA:

- Análisis de Ítems.
- Análisis Factorial de los ochenta ítems del Cuestionario.
- Análisis Factorial de los cuatro Factores Teóricos (veinte ítems cada uno).
- Análisis Factorial de las Medidas obtenidas en cada uno de los Estilos.
- Análisis Comparativo entre ellos.

Para contrastar las dieciocho hipótesis planteadas:

- Análisis de Varianza simple.
- Pruebas de contraste de Sheffé.
- Análisis Discriminante.

Conclusiones

- El estudio y la aplicación en el aula de los Estilos de Aprendizaje es uno de los caminos más fecundos que conocemos para individualizar la instrucción, da su fundamento científico.
- Existen múltiples Instrumentos de diagnóstico de los Estilos de Aprendizaje, dirigidos a colectivos diversos y con metodologías muy distintas. La Teoría de los Estilos de Aprendizaje ha sido aplicada con éxito en todos los niveles educativos.
- El ajuste de Estilo de aprender y Estilo de enseñar se correlaciona positivamente con el éxito académico.
- Nuestro Cuestionario CHAEA ofrece una Fiabilidad aceptable y Validez comprobada para ser aplicado a universitarios españoles.
- Hay diferencias significativas en los cuatro Estilos de Aprendizaje entre los universitarios según la Facultad en que estudian y según la Facultad que eligieron en primer lugar.
- No hay diferencias significativas en ninguno de los cuatro Estilos de Aprendizaje respecto a las profesiones del padre o de la madre, el número de hermanos, la ciudad de origen.
- En las restantes hipótesis existen diferencias significativas en alguno/s de cuatro Estilos de Aprendizaje.
- Los Estilos de Aprendizaje pueden mejorarse con «tratamientos» específicos de optimización.
- El enfoque prospectivo ante las Oportunidades de Aprendizaje y el Diario Personal de Aprendizaje son dos formas prácticas de «vivir aprendiendo».

Autor: María Frontera Sancho

Dirección: Rubén Darío, 7, 0-3
50012 Zaragoza

Director: Arturo de la Orden Hoz
Dpto.: Métodos de Investigación y Diagnóstico en Educación
Dirección: General Yagüe, 11, 8º E
28020 Madrid

Descriptores

Pensamiento matemático. Resolución de problemas aritméticos de enunciado verbal. Conceptos de adición y sustracción. Estrategias de solución. Errores en el proceso de solución.

Bibliografía fundamental

- BAROODY, A. J.: *El pensamiento matemático de los niños*, Aprendizaje visor y Centro de publicaciones del MEC, Madrid, 1988.
- BAROODY, A. J. y GINSBURG, H. P.: «The relationship between initial meaningful and mechanical knowledge: of arithmetic». En J. HIEBERT (Ed.): *Conceptual and procedural knowledge: The case of mathematics*, Hillsdale, N. J.: Lawrence Erlbaum Associates, 1986, pp. 75-112.
- BRIARS, D. J. y LARKIN, J. H.: «Ab integrated model of skills in solving elementary word problems» *Cognition and Instruction*, 1, 1984, 245-296.
- CARPENTER, T. P.: «Conceptual knowledge: Implications from research on the initial learning of arithmetic». En J. Hiebert (ed.): *Conceptual and procedural knowledge: The case of mathematics*, 1986, 113-132. Lawrence Erlbaum Associates. Hillsdale.
- CARPENTER, T. P.; HIEBERT, J. y MOSER, J. M.: «Problem structure and first-grade children's initial solution processes for simple addition and subtraction problems». *Journal for Research in Mathematics Education*, 1981, 12, 27-39.
- CARPENTER, T. P. y MOSER, J. M.: «The development of addition and subtraction problem solving skills. En T. P. Carpenter; J. M. Moser y T. A. Romberg (Eds.): *Addition and Subtraction: a cognitive perspective*. Erlbaum. Hillsdale, N. J., 1982.
- DE CORTE, E.; VERSCHAFFEL, L.: «The effect of semantic structure on first graders' strategies for solving addition and subtraction word problems». *Journal for Research in Mathematics Education*, 1987, 18, 363-381.
- HIEBERT, J. y LEFEVRE, P.: «Conceptual and Procedural Knowledge in Mathematics: An Introductory Analysis». En Hiebert (Ed.): *Conceptual and Procedural Knowledge: The case of Mathematics*. Lawrence Erlbaum Associates. Hillsdale, 1986.
- IBARRA, C. G. y LINDVALL, C. M.: «Factors associated with the ability of kindergarten children to solve simple arithmetic story problems. *Journal of Educational Research*, 1982, 75, 149-155.
- RILEY, M. S.; GREENO, J. G. y HELLER, J. I.: «Development of children's problem-solving ability in arithmetic». En H. Ginsburg (Ed.): *The development of mathematical thinking*. Academic Press. Nueva York, 1982.

Problema investigación

La investigación se centra en el estudio del razonamiento matemático de los niños, su habilidad para resolver problemas matemáticos elementales, las estrategias

Autor: María Frontera Sancho

ADQUISICIÓN DE LOS CONOCIMIENTOS MATEMÁTICOS
BÁSICOS. UNA PERSPECTIVA COGNITIVA

CAT ⁽¹⁾	ANO ⁽²⁾	CLASIFICACIÓN ⁽³⁾
D	1 9 9 1	6 1 0 4 0 1
		N.º CITAS ⁽⁴⁾ N.º PÁGINAS
		8 3 6

que utilizan y los errores que cometen en su solución antes de recibir instrucción formal, así como los cambios que tienen lugar durante los dos primeros años de la escolaridad. Es decir, se trata de analizar la transición del conocimiento matemático espontáneo al conocimiento formal. El estudio va encaminado a aportar algún conocimiento sobre las bases para una didáctica de las matemáticas en Preescolar y primeros años escolares.

Muestra y método de muestreo

La muestra está formada por 48 niños, 24 de 1º y 2º de Preescolar y 24 de primer ciclo de EGB, repartidos por igual en los cuatro cursos y seleccionados según su nivel de rendimiento académico valorado por el profesor, de modo que cada curso está representado por 4 alumnos de nivel alto, 4 de nivel medio y 4 de nivel bajo.

Metodología del trabajo

Se trata de un estudio esencialmente cualitativo. La metodología básica consistió en entrevistas clínicas practicadas individualmente a los niños, en las que se les planeó tareas matemáticas.

El trabajo se centra en el estudio de la resolución de problemas aritméticos sencillos de enunciado verbal y así, presentamos a los niños en días sucesivos un total de 17 tipos distintos de problemas, siguiendo la clasificación de Carpenter y Moser (1982), atendiendo a la estructura semántica de los mismos.

La misión del resto de las tareas fue la de permitir una valoración más matizada y profunda de la habilidad de los niños para resolver problemas. Este fue el caso de las pruebas piagetianas de Conservación del número, Seriación e Inclusión de clases; de las que evalúan la comprensión de las propiedades aritméticas de Inferencia de la operación realizada, Inversión, Compensación, Asociatividad y Conmutatividad, así como de la encuesta sobre el conocimiento de la operatividad aditiva.

Técnicas de análisis

Frecuencias y porcentajes de respuestas correctas, tipos de estrategias utilizadas y de errores cometidos por los niños en cada una de las tareas, distinguiendo cursos y niveles de rendimiento.

Para el estudio comparativo y el estudio de relaciones, las principales técnicas estadísticas utilizadas han sido: la prueba ji-cuadrado, el Coeficiente de contingencia, la prueba de la probabilidad exacta de Fisher, la t de Student y la razón crítica Z de MacNemar.

Conclusiones

Además de las conclusiones relativas a la dificultad y a la naturaleza de las estrategias y errores que aparecen en los distintos tipos de problemas en función del curso y el nivel de rendimiento de los niños (lo que podría contribuir a guiar la planificación de la enseñanza de las matemáticas en los primeros cursos), deseamos destacar algunas que llaman especialmente la atención:

La escolaridad no sólo no contribuye en todos los casos a capacitar a los niños para una mejor resolución de los problemas sino que, en determinadas circunstancias supone un obstáculo.

La habilidad para resolver problemas es relativamente independiente del curso al que hayan llegado los niños dentro del plan de estudios de matemáticas.

Un avance en las estrategias utilizadas no lleva consigo necesariamente una mejor ejecución en los problemas.

Se pone de manifiesto, por una parte, que para resolver problemas sencillos de suma y resta, no es necesario el conocimiento de estas operaciones aritméticas y, por otra, que tal conocimiento no es suficiente para hallar la solución si no se comprende su estructura semántica.

ESTUDIO PSICOANALÍTICO DE CUENTOS
INFANTILES

Autor: Gerardo Gutiérrez Sánchez

Autor: Gerardo Gutiérrez Sánchez

Dirección: Vereda de Ganapanes, nº 35
28035 Madrid

Director: Ángel Lázaro Martínez
Dpto.: Métodos de Investigación y Diagnóstico en Educación
Dirección: Ríos Rosas, 51, 6º G
28003 Madrid

Centro: Facultad de Educación de la U.C.M.

Descriptor

Cuentos y romances populares; Incesto y seducción traumática; Mutilación y castración simbólica; Inconsciente del texto; Psicoanálisis; Interpretación; Asociación libre; Escuela analítica; Transferencia; Significado y significante; Imaginario, Simbolismo y Real.

Bibliografía fundamental

BELLEMIN-NOEL, J. (1989): *Psychanalyse et Littérature*, Presses Universitaires de France, París.
 BETTELHEIM, B. (1977): *Psicoanálisis de los cuentos de hadas*, Crítica, Barcelona.
 DOR, J. (1991): *El padre y su función en psicoanálisis*, Nueva Visión, Buenos Aires.
 FREUD, S. (1976): *Obras Completas*, Amorrortu, Buenos Aires.
 GEORGIN, R. (1988): *De Lévi-Strauss a Lacan*, Nueva Visión, Buenos Aires.
 LAPLANCHE, J. (1988): *Castración. Simbolizaciones. Problemáticas II*, Amorrortu, Buenos Aires.
 MANNONI, O. (1979): *La otra escena. Claves de lo imaginario*, Amorrortu, Buenos Aires.
 MILLER, J. A. (1986): *Recorrido de Lacan*, Manantial, Buenos Aires.
 MILLOT, C. (1982): *Freud Antipedagogo*, Paidós, Barcelona.
 RODRÍGUEZ ALMODOVAR, A. (1989): *Los cuentos populares o la tentativa de un texto infinito*, Secretariado de Publicaciones de la Universidad de Murcia, Murcia.

Problema investigación

Se parte de la hipótesis de que los cuentos populares, que han pervivido durante siglos y en las más distintas culturas, tienen alguna relación con aquello que constituye, desde una óptica psicoanalítica, la estructura psicológica del ser humano.

Se propone la interpretación psicoanalítica como una forma de acercamiento al sentido de los cuentos. Para ello, es preciso solventar previamente importantes problemas teóricos: ¿bajo qué condiciones la interpretación psicoanalítica, que nace formando parte de una metodología clínico-terapéutica, es aplicable al análisis de textos?

Dos hipótesis por tanto: a) los cuentos, los romances constituyen historizaciones imaginarias de elementos que pertenecen a la estructura psíquica. b) la interpretación psicoanalítica de los cuentos es válida, en determinadas condiciones, para acceder a esos elementos de la estructura.

CAT ⁽¹⁾		AÑO ⁽²⁾		CLASIFICACIÓN ⁽³⁾	
D	1	9	9	6	1
	9	2		1	0
				3	0
				6	
N.º CITAS ⁽⁴⁾		N.º PÁGINAS			
4	0	7	8	5	3

Muestra y método de muestreo

Sobre una población de dos mil cuentos populares, se realiza una selección (siguiendo el criterio de accesibilidad del lector medio a las publicaciones, combinado con el de rigor y calidad en las versiones: editoriales muy acreditadas en literatura infantil y publicaciones del CSIC) que incluye:

12 versiones del tipo «La niña sin brazos».

69 versiones del tipo del ciclo «La niña perseguida».

21 versiones literarias antiguas (desde el s. XII) de «La niña sin brazos».

Metodología del trabajo

Recogida de información y análisis de fuentes: la investigación se realiza con material directo y seleccionado, y se consideran gratuitas y faltas de rigor aquellas conclusiones extraídas por autores que no han tenido acceso a los tipos originales y a las múltiples versiones de cada uno de ellos.

Comparativa: se confeccionan cuadros de cada tipo estudiado cuya comparación permite extraer alguna conclusión significativa.

Psicoanalítica: única que puede ser potencialmente eficaz para acceder al inconsciente del texto. Concepción ésta, del inconsciente, que está lejos de entenderse como el reservorio de los significados del sujeto o de la cultura. Más bien el inconsciente del que hablamos —cuando lo relacionamos con el texto—, tendría que ver con los significantes fundamentales cuya presencia o ausencia conforman la estructura simbólica de uno y de otra (sujeto y cultura).

Técnicas de análisis

La interpretación psicoanalítica. Se proponen unas «condiciones de posibilidad» para la interpretación: 1) que se de la *asociación libre*; 2) que el intérprete mantenga una actitud de *escucha analítica*; y 3) que entre texto y lector se establezca un *vínculo de tipo transferencial*.

Se ha considerado como equivalente a 1) presencia del tema en otras versiones cuentísticas; en otras formas de lo popular (romances); en versiones cultas con variantes altamente significativas; en mitos, tragedias, etc. Equivalente a 2) realizar una lectura sensible a los sinsentidos, a las elisiones, a los añadidos, a los sobreentendidos, etc., es decir, que vaya marcando dónde queda rota la continuidad semántica del relato. Resaltando los enigmas del cuento. Equivalente a 3) precisamente esa ruptura del sentido que «no está en el texto si no hay alguien que la escuche», y que «no puede ser escuchada si en cierto modo no se produce en el texto».

Conclusiones

En el tipo «La niña sin brazos» tenemos el siguiente esquema: 1) deseo incestuoso-mutilación-destierro-encuentro con príncipe que la desposa; 2) nuevo destierro decretado por la madre-recuperación milagrosa de las manos; 3) reencuentro y rehabilitación. Sólo hemos investigado la primera parte.

Eso nos ha llevado a estudiar los siguientes conceptos: la *seducción traumática paterna* (deseo incestuoso), la *castración simbólica* (mutilación), la *desfloración ritual* (seducción paterna + castración simbólica), concluyendo que:

Se trata de «fantasmas originarios» que están en la base y estructuran toda la actividad fantasmática (de pensamiento, de deseo) del sujeto: la fantasía de *seducción paterna* es la versión imaginaria de la operación que hemos situado en el *origen de la sexualidad del sujeto*. La *mutilación-castración simbólica* representa el *origen de la diferencia, de la falta y del deseo*. La *desfloración ritual* es la versión imaginaria del *origen de la mujer como sujeto social sexuado y deseante*.

Este saber simbólico se transmite con la transmisión del cuento.

**FACTORES DE EFICACIA DE LOS CENTROS ASOCIADOS
DE LA UNED**

Autor: Catalina Martínez Mediano

Autor: Catalina Martínez Mediano

Dirección: Dr. García Tapia, nº 114. 8º A
28030 Madrid

Director: Ramón Pérez Juste

Dpto.: Métodos de Investigación y Diagnóstico en Educación

Dirección: Senda del Rey s/n
28040 Madrid

Descriptor

Investigación evaluativa; Evaluación de Centros; Centros Asociados de la UNED.

Bibliografía fundamental

DE LA ORDEN, A. (Ed.) (1985): «Investigación evaluativa». *Diccionario de Ciencias de la Educación*. Anaya, Madrid.

EVERTSON, C. M. y GREEN, J. L. (1986): «Observation as inquiry and method». En M. C. WITTRICK (Ed.): O. c.

FUENTES VICENTE, A. (1986): *Procesos funcionales de eficacia de la escuela. Un modelo causal*. Tesis Doctoral. Facultad de F³ y CC.EE. Universidad Complutense, Madrid.

PÉREZ JUSTE, R. y MARTÍNEZ ARAGÓN, L. (1989): *Evaluación de centros y calidad educativa*. Cincel, Madrid.

PÉREZ JUSTE, R. y otros: *Estudios sobre el curso de acceso a la UNED para mayores de 25 años*. UNED. E.E.D. 15, Madrid. (En prensa).

ROSENSHINE, B. y STEVENS, R. (1986): «Teaching functions». En M. C. WITTRICK (Ed.): O. c.

SCHEERENS, J. (1990): «School effectiveness research and the development of process indicator of school functioning». *School effectiveness and School Improvement*, 1, 1, (61-80).

STUFFLEBEAM, D. L. y SHINKFIELD, A. J. (1987): *Evaluación sistemática*. Paidós-MEC, Barcelona, Buenos Aires.

WITTRICK, M. C.: *Handbook of research on teaching*. Third Ed. MacMillan P. Co, New York.

WORTHEN, B. R. y SANDERS, J. R.: *Educational Evaluation*. Longman, New York, London.

Problema investigación

Se trataba de estudiar la influencia que las instituciones docentes tienen en los resultados de enseñanza de sus alumnos, a través de los elementos que las integran (sus edificios, sus recursos económicos, sus organizaciones, las funciones docentes y la satisfacción de sus miembros).

Se hipotetiza que en los Centros Asociados de rendimiento académico alto en el Curso de Acceso Directo aparecen en mayor medida que en los Centros Asociados de rendimiento académico bajo en el mismo curso variables predictoras de eficacia. (Las variables se citan en el apartado Metodología).

Con este trabajo se pretendía, además, contribuir a la línea de investigación del Departamento MIDE de la UNED sobre «Análisis y valoración del modelo español de educación superior a distancia» y al encargo de la Junta de Gobierno de la UNED de estudiar el curso de acceso directo.

CAT ⁽³⁾	ANO ⁽²⁾	CLASIFICACIÓN ⁽³⁾
D	1 9 9 1	5 8 0 1 0 5

N.º CITAS ⁽⁴⁾	N.º PÁGINAS
1 6 6	3 8 0

Muestra y método de muestreo

La selección de los Centros Asociados se realizó tras un análisis de las calificaciones obtenidas por los alumnos del Curso de Acceso Directo (CAD) durante los años académicos de 1985-86 a 1988-89, por Centros Asociados.

Se eligieron seis Centros Asociados de rendimiento académico extremo en el CAD. Dos fueron analizados en el estudio piloto, en 1988-89 y cuatro en el estudio definitivo en el año académico siguiente, 1989-90.

Metodología del trabajo

La metodología utilizada es la propia de la investigación evaluativa con un enfoque holístico que pretende abordar la realidad de los centros en el marco en que se inscriben. Se han utilizado técnicas propias de la investigación cualitativa (observación, entrevistas y lecturas de documentos) y de la cuantitativa (pruebas de conocimientos previos y cuestionarios de opinión). El trabajo de campo fue realizado en dos momentos en cada uno de los cuatro Centros Asociados estudiados: al comienzo del curso y durante el segundo trimestre (1989-90).

Fue necesario elaborar instrumentos específicos sobre las variables de estudio (Situación y recursos del Centro Asociado, Funciones de la dirección, Organización académica, Criterios y actividad docente del profesor tutor, Características profesionales y académicas del alumno y Satisfacción de sus miembros).

Técnicas de análisis

El seguimiento del rendimiento académico de los alumnos del Curso de Acceso (CAD) mostró una distribución de Centros Asociados con una amplitud de 3,55 desviaciones típicas. Este hecho centró el problema de estudio y sugirió la conveniencia de identificar los factores de eficacia de los Centros implicados en los niveles de éxito académico de los alumnos del CAD en ellos inscritos y contrastar la información recogida en cada Centro.

Por ello, para el análisis de datos se han utilizado técnicas de contraste, mediante la prueba de «chi» cuadrado, entre la información procedente de los diferentes centros estudiados y triangulación entre las diferentes fuentes de información (la dirección y el personal del Centro, los profesores tutores, los alumnos y el observador).

Conclusiones

Las conclusiones no pueden desligarse de las peculiaridades de los Centros Asociados (CCAA) de la UNED, del Curso de Acceso o de la modalidad de enseñanza a distancia, por lo que las diferencias encontradas entre los CCAA de alto y bajo rendimiento no muestran un perfil nítido como harían esperar las diferencias iniciales en rendimiento. De este modo, algunas hipótesis sólo quedan confirmadas parcialmente y en otras no puede rechazarse la hipótesis nula.

No obstante, a partir del contraste de hipótesis hemos encontrado suficientes factores que pueden explicar las características diferenciales de los centros de rendimiento académico alto. Se observan estas diferencias en la mejor ubicación de sus edificios, financiación y recursos. Se da un seguimiento más continuado de la organización del centro por la dirección. Ofrecen más horas de atención académica a los alumnos y predominan criterios y funciones docentes más prácticos y realistas, observándose mayor satisfacción del profesor tutor y de los alumnos con los criterios organizativos del centro.

Autor: María Antonia Naranjo Díaz

Dirección: Cardenal Herrera Oria, 3, 1º
04005 Almería

Director: Arturo de la Orden Hoz

Dpto.: MIDE

Dirección: Ciudad Universitaria
28040 Madrid

Centro: Facultad de Educación. Universidad Complutense

Descriptores

Estilo cognitivo dependencia-independencia de campo; Diferenciación Psicológica; Preferencias profesionales; Orientación académico-profesional; Madurez vocacional; Conducta vocacional.

Bibliografía fundamental

CARRETERO, M. (1980): «Investigaciones sobre el pensamiento personal», *Revista de Psicología General y Aplicada*, 35.

DE LA ORDEN, A. y GARCÍA RAMOS, J. M. (1985): «Estilo cognitivo y Orientación Académico Profesional en la Universidad». IX Congreso Internacional de Educación y trabajo. Madrid.

HOLLAND, J. L. (1981): «Técnicas de elección vocacional». *Tipos de personalidad y modelos ambientales*. Trillas, México.

HUTEAU, M. (1981): *Cognition y personalidad. La dependence-independance du champ*. Thèse. Laboratory de Psychologie Differentielle de l'Université René Descartes. Paris.

OSIPOW, S. H. (1969): «Cognitive styles and educational-vocational preferences and selection», *Journal of counseling Psychology*, 16.

PASCUAL-LEONE, J. (1974): «A neo-piagetian process-structural model of witkin's psychological differentiation». Second International Conference of the International Association for Cross-Cultural Psychological. Ontario.

SUPER, D. E. (1969): «Vocational developmental theory: persons, positions and processes», *Consulting Psychology*, 1, pp. 2-9.

WITKIN y COL. (1977): «Role of the field-dependent and field independent cognitive styles in academic-evolution: a longitudinal study», *Journal of Education Psychology*, 69 (2), pp. 197-211.

WITKIN, H. A. y GOODENOUGH, D. R. (1985): *Estilos cognitivos. Naturaleza y orígenes*. Pirámide, Madrid.

Problema investigación

Nuestro planteamiento persigue los siguientes objetivos:

1. Estudiar un conjunto de factores aptitudinales y circunstanciales a fin de determinar si tales factores son o no concurrentes significativamente en la elección académico-profesional que hacen los alumnos de COU.
2. De todos los factores queremos destacar el estilo cognitivo DIC y los intereses profesionales, así como su grado de influencia e implicación en la elección profesional.
3. Comprobar qué variables serían discriminantes.
4. Conocer la percepción que de la Orientación y su necesidad manifiestan los estudiantes.

El objetivo pedagógico final es proponer nuevos factores en el currículum de secundaria y elaborar un método de intervención psicopedagógica en la Orientación Académico-Profesional.

Se redactan cuatro hipótesis, tales como que:

- a) Los aspectos del funcionamiento diferencial cognitivo y capacidades individuales como: estilo cognitivo DIC, inteligencia e intereses profesionales.
- b) Sexo.

ESTILO COGNITIVO, INTERESES Y ELECCIÓN
ACADÉMICO-PROFESIONAL EN LA ORIENTACIÓN
DE LOS ALUMNOS DE COU

Autor: María Antonia Naranjo Díaz

CAT ⁽¹⁾	ANO ⁽²⁾	CLASIFICACIÓN ⁽³⁾
D	1 9 9 2	6 1 0 4 0 1
N.º CITAS ⁽⁴⁾ N.º PÁGINAS		
	5 7 9	5 1 4

- c) Los factores circunstanciales: estudios del padre, tener o no beca, el número de hermanos, el medio geográfico, el régimen educativo e irse a estudiar fuera de Almería.
- d) Los factores académicos: opción de COU, el historial académico, el rendimiento escolar y sus gustos.
Todos estos variables pueden estar relacionados con la elección de carrera.

Muestra y método de muestreo

La muestra recogida pertenecía a una población de 564 alumnos de COU (opciones de Letras, Ciencias y Mixto) en la provincia de Almería, en el curso 87-88, de ambos sexos y de toda la provincia.

El muestreo fue aleatorio, al azar, los pueblos con instituto, del norte, sur, este y poniente almeriense, además de estratificado, según el medio geográfico (rural, urbano) y el tipo de Centro (público y privado).

Metodología del trabajo

El esquema general de la investigación expone los factores supuestamente implicados en la elección de carrera, fruto de la madurez y conducta vocacional del alumno.

Los instrumentos de medida utilizados fueron de dos tipos:

- a) estandarizados: test de inteligencia general D-70, de Rennes y Kuzosvski, test de figuras enmascaradas, colectiva de Witkin y de intereses profesionales de Kuder.
- b) encuesta elaborada, aplicada y tabulada para este trabajo.

El resto de la información procedía:

- a) de las actas finales de evaluación de COU de cada Centro, las calificaciones obtenidas por cada estudiante.
- b) La información recogida en los distintos ítems de la encuesta ya citada.

Con toda esta información se realizó una tabla general con datos numéricos para su procesamiento informático.

Técnicas de análisis

El estudio se configura como un diseño ex-post-facto que quiere probar la validez diferencial y predictiva de cada variable.

El programa estadístico utilizado está integrado en el paquete B M D P, implementado en el Departamento de Estadística e Investigación Operativa de la Universidad de Granada.

Este análisis ha sido doble:

- 1) Análisis discriminante para diferenciar grupos de elección de carrera en un colectivo de variables.
- 2) Estudio de X^2 para constatar la relevancia del estilo DIC, ver las relaciones significativas entre la DIC y la elección de carrera, y entre la DIC y las variables clasificadas por el análisis discriminante.

Por último se realizó un análisis de frecuencias de porcentajes de las opiniones de los alumnos recogidas en el cuestionario.

Conclusiones

En la elección de carrera que hacen los alumnos de COU se manifiestan significativas e influyentes los siguientes factores:

1. *El rendimiento académico* en las distintas disciplinas académicas. Si bien la materia de LENGUA se manifiesta como la más significativa de todos (ya que implica capacidades como la aptitud, comprensión y fluidez verbal, así como ser el mayor código de comunicación) a partir de la cual se elaboran los aprendizajes significativos.

Este factor, es pues, el principal indicador de la conducta vocacional.

2. *La opción de COU* elegida, ya que dibuja o perfila los campos profesionales.

3. *Los gustos*. Preferencias profesionales personales. El 90% de los alumnos dicen que eligen esa carrera «porque me gusta». Aun siendo importante este factor por lo que afecta a la motivación, no deja de evidenciarse una falta de la actividad orientadora en el sistema educativo.

4. *El medio geográfico* en el que vive el alumno influye, debido a que determina el contexto cultural, así como las pautas sociales y culturales.

5. *El tener o no tener beca* determina igualmente la continuación de los estudios superiores. Conviene destacar que aunque el sexo no influye en la elección de carrera, si se evidencia su relación con el estilo cognitivo DIC, es decir, los varones tienden hacia las carreras de ciencias y son más independientes de campo; las mujeres tienden a las carreras de letras y son más dependientes del campo cognitivo.

Como resultado y conclusión pedagógica se propone un modelo de intervención psicopedagógica en la Enseñanza Media o Secundaria que recoge estos factores citados así como las bases psicopedagógicas adecuadas.

Autor: Ana María Porto Castro

Dirección: Juan Bautista Andrade, 50, 4º A
36005 Pontevedra

Director: José Cajide Val

Dpto.: Métodos de Investigación y Diagnóstico en Educación

Dirección: F. de Fª y CC. de La Educación, Campus Universitario
Santiago de Compostela

Centro: F. de Fª y CC. Educación. Sección CC. de la Educación

Descriptorios

Identificación del alumno muy capacitado. Programas para el estudiante muy capacitado: desarrollo y elaboración. Diseño y evaluación de un programa en el área de matemáticas. Estudio experimental. Pruebas de significación.

Bibliografía fundamental

- ARNAU, G. J. (1984): *Diseños experimentales en psicología y educación*, Trillas, México.
- CORIAT, A. (1990): *Los niños superdotados*, Herder, Barcelona.
- GALLAGHER, J. J. (1985): *Teaching the gifted child*, Allyn and Bacon, Inc., London.
- HOROWITZ, F. D. y O'BRIEN, M. (1985) (eds.): *The gifted and talented: Development perspectives*, American psychological Association, Washington, D.C.
- HUTTEMA, B. E. (1980): *The analysis of covariance alternatives*, John Wiley, New York.
- PASSOW, A. H. (1984): «Educating of the gifted», *Prospects*, 14, 2; 2.045-2.056
- RADFORD, J. (1990): *Child prodigies and exceptional early achievers*, Harvester Wheatsheaf, London.
- RENZULLI, J. S. (1986): *Systems and models for developing programs for the gifted and talented*, Creative Learning Press, Inc., Mansfield Center Connecticut.
- STRAKER, A. (1983): *Mathematics for gifted pupils*, Schools Council Publication, London.
- WHEATLEY, G. H. (1983): «A mathematics curriculum for the gifted and talented», *Gifted Child Quarterly*, 27, 2, 77-80.

Problema investigación

Estudio de la problemática del niño muy capacitado abordando tres grandes áreas fundamentales. Su identificación, el diseño y desarrollo de programas y la evaluación de los mismos. Los puntos más relevantes de estos aspectos se concretan en: ¿cómo identificarlos?, es decir, en el proceso de medida y selección seguido; ¿cómo enseñarlos?, esto es, las diversas alternativas que se han presentado y la concreción en un programa en el área de matemáticas, así como en la evaluación de sus efectos, todo ello con la finalidad de recoger aquellos aspectos más significativos que puedan ayudar a sentar las bases de una acción educativa en su favor.

En base a ello la investigación se centra en unos objetivos fundamentales:

- En primer lugar, comprobar las líneas de investigación y problemas relacionados con la identificación de estos niños, los métodos y procedimientos que den prueba y permitan una adecuada evaluación de las capacidades y potencial que presentan.
- Examinar las principales medidas educativas desarrolladas y los elementos que sientan las bases para el desarrollo de programas y el diseño del currículum adaptado a sus características y necesidades.

EFFECTOS DE UN PROGRAMA EN EL ÁREA DE MATEMÁTICAS
CON ALUMNOS MUY CAPACITADOS

Autor: Ana María Porto Castro

CAT^(a)

D	1	9	9	1															
---	---	---	---	---	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

ANO^(a)

CLASIFICACIÓN^(a)

N.º CITAS^(a) N.º PAGINAS

	2	9	5		4	6	5
--	---	---	---	--	---	---	---

- Llevar a cabo un proceso de selección para localizar a los alumnos que puedan ser considerados como muy capacitados y, finalmente, afrontar aquellos aspectos fundamentales de un programa centrado en el área de matemáticas, ponerlo en práctica y evaluar sus efectos.

Muestra y método de muestreo

Un total de 204 alumnos de 12-14 años, pertenecientes a 6 colegios públicos de E.G.B. ubicados en núcleos de población urbanos, semiurbanos y rurales sirven de base para el estudio e integrar 3 muestras principales:

- una muestra de 18 alumnos previamente seleccionados como muy capacitados, dividida en 2 grupos, el experimental E_1 ($n=9$) y el control C_1 ($n=9$);
- una muestra de 186 alumnos no seleccionados como muy capacitados, dividida en 2 grupos, el experimental E_2 ($n=85$) y el control C_2 ($n=101$); una muestra al azar de alumnos no seleccionados como muy capacitados dividida en 2 grupos, el experimental E_3 ($n=9$) y el control C_3 ($n=9$).

Metodología del trabajo

El estudio se divide en dos partes. Desarrolladas las bases teóricas del mismo, en la segunda parte se procede inicialmente a la medida y selección de un grupo de alumnos de 7º curso de E.G.B. en base a los resultados obtenidos en distintas pruebas psicométricas estandarizadas, el rendimiento escolar y las nominaciones de profesores y compañeros.

Una vez seleccionados los alumnos y comprobada, a partir del tratamiento estadístico de los datos e información recogida, la superioridad de éstos en las medidas efectuadas, se elaboró un programa de matemáticas centrado en la solución de problemas procediendo, finalmente, a la evaluación de los resultados y efectos del mismo.

Técnicas de análisis

Los datos e información obtenidos durante el proceso de medida y selección de los alumnos muy capacitados fueron analizados mediante:

- Prueba «t» para muestras independientes. ANOVA Unidireccional para la comprobación de la Homogeneidad de las Varianzas. Prueba de Kolmogorov-Smirnov.

En la evaluación y estudio experimental de los efectos del programa se utilizaron las siguientes técnicas de análisis:

- Prueba «t» para dos muestras independientes. Cálculo de la Razón Crítica Z. ANOVA. Prueba de Scheffé y Prueba de Tukey para las Comparaciones Múltiples: ANCOVA. Análisis de Regresión Múltiple.

Conclusiones

Los alumnos seleccionados como muy capacitados se diferencian y presentan, en términos de grupo, una destacable superioridad con respecto a sus compañeros no seleccionados, en todas las variables medidas y sometidas a estudio.

El programa puesto en práctica es efectivo en los grupos experimentales al compararlos con sus respectivos grupos de control. Estos efectos son especialmente favorables para los alumnos seleccionados quienes alcanzan unos resultados superiores a sus compañeros igualmente muy capacitados que no participan en el mismo.

La superioridad que muestran los alumnos seleccionados en su potencial intelectual, las aptitudes escolares verbales, numéricas y de razonamiento, el dominio de las matemáticas que poseen, y muy especialmente, en su rendimiento escolar anterior son predictores efectivos de los resultados o efectos del programa.

El principal resultado del estudio es que los alumnos muy capacitados, cuando trabajan con un programa ajustado a sus capacidades se diferencian de sus compañeros también muy capacitados que siguen una enseñanza normal. De ello se deriva la necesidad de elaborar programas en los que se contemple una atención especial a estos alumnos para que puedan aprovechar sus capacidades y potencial al máximo y alcancen un desarrollo pleno acorde con sus posibilidades.

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 4.000 ptas.

A.I.D.I.P.E.
BOLETÍN ASOCIADO

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

N.º de cuenta N.º de libreta

N.º Agencia Domicilio Agencia

Población C.P.

(Firma)

A.I.D.I.P.E.
BOLETÍN DOMICILIACIÓN BANCARIA

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

N.º de cuenta N.º de libreta

N.º Agencia Domicilio Agencia

Población C.P.

(Fecha y Firma)

