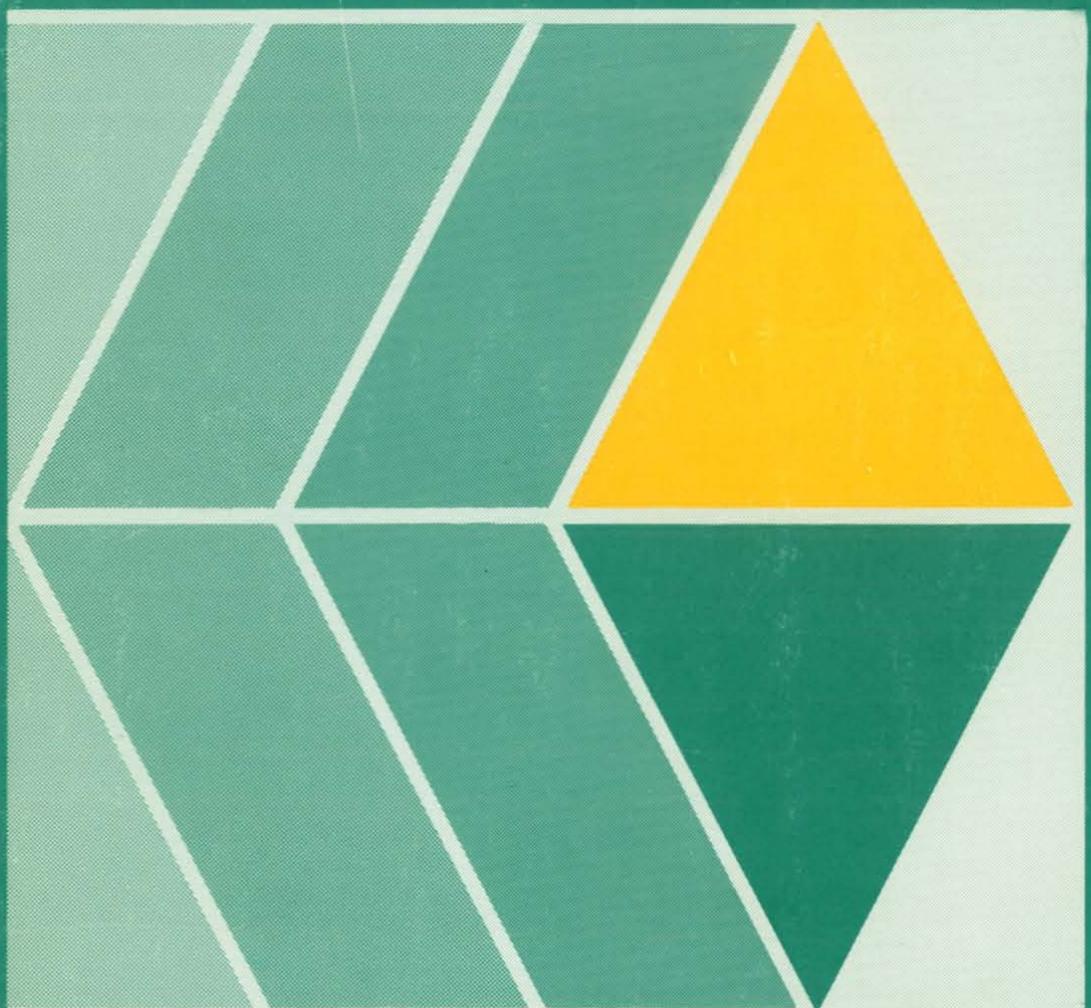


REVISTA  
INVESTIGACION  
EDUCATIVA



Vol. 3, N.º 6, 2.º semestre 1985

N.º EXTRAORDINARIO



NUEVAS TENDENCIAS DE LA INVESTIGACION  
SOBRE RENDIMIENTO ACADEMICO



ASOCIACIÓN INTERUNIVERSITARIA  
DE INVESTIGACIÓN PEDAGÓGICA  
EXPERIMENTAL

REVISTA INVESTIGACION EDUCATIVA

Vol. 3, N.º 6, 2.º semestre 1985

# Colección ESTADÍSTICA

Libros de Estadística aplicada a las Ciencias Humanas y más en concreto al campo de la Educación.

La experiencia docente de sus autores, como profesores de esta materia en Facultades universitarias, se deja traslucir a lo largo de estas obras, pensadas para alumnos de Escuelas de Formación del profesorado y Facultades de Pedagogía y Psicología.

## NORMATIVA PARA LOS COLABORADORES

El objetivo de la *Revista de Investigación Educativa* es promover el intercambio de información acerca de investigaciones empíricas de carácter educativo. Consta de cuatro secciones: a) TRABAJOS de investigación; b) LINEAS de investigación; c) NOTICIAS y d) FICHAS-RESUMEN. Todo profesional que desee colaborar en la revista deberá atenerse a dichas secciones y a las siguientes indicaciones:

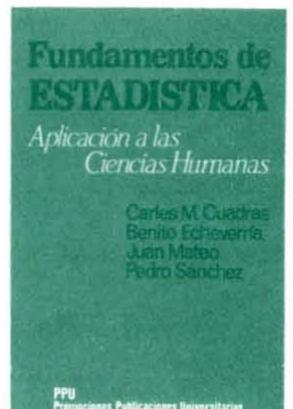
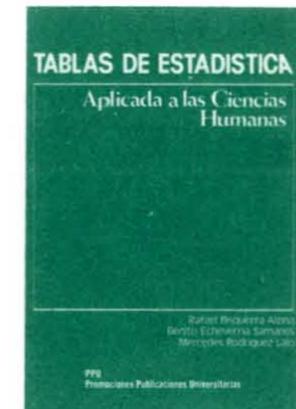
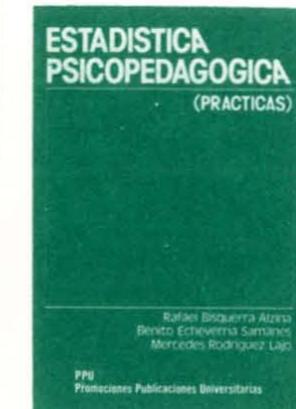
1.— Para los "Trabajos" y la "Linea" se enviarán original y copia mecanografiadas en DIN A4 a doble espacio y numeradas. No puede exceder de 10 hojas en el caso de "Trabajos", ni de 20 en la "Linea". En hoja aparte se recogerá el sumario (Máximo 20 líneas), los principales descriptores y una suscinta biografía del autor, así como la dirección completa de su lugar de trabajo.

Las referencias bibliográficas se presentarán al final por orden alfabético (APELLIDO, Inicial del nombre, Año, —señalando "a", "b", "c"... cuando coinciden varias en el mismo—. Título, Editorial, Ciudad). Las llamadas a estas referencias se harán dentro del artículo (Apellido, Inicial del nombre, Año, "a", "b"... etc). Se aconseja insertar las notas también al final.

2.— Las "Noticias" no pasaran de 2 DIN A4, mecanografiadas a doble espacio. Se ha de enviar original y copia, así como material gráfico, si se desea.

3.— Las "Fichas" se cumplimentarán según los modelos propuestos e impresos por la misma revista.

Los números se cierran los días 1 de Mayo y 1 de Noviembre de cada año.

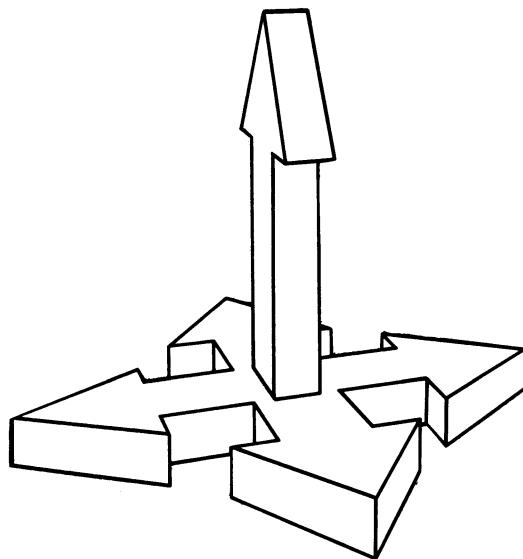


**ESTADÍSTICA PSICOPEDAGÓGICA:** La obra recoge esquemáticamente los conceptos fundamentales de *Estadística Descriptiva* y presenta gran número de ejercicios (resueltos, sin resolver y con soluciones), así como pruebas de autovaloración para el estudiante (P.V.P.: 1.500 pts.).

**TABLAS DE ESTADÍSTICA:** Compendio de las tablas más usuales en investigación educativa con un resumen de las principales fórmulas de Estadística Descriptiva (P.V.P. 450 pts.).

**FUNDAMENTOS DE ESTADÍSTICA:** Obra de *Estadística Inferencial* con una cuidada fundamentación probabilística y un exhaustivo temario de las principales técnicas estadísticas, al que acompaña numerosos ejercicios relacionados con el campo de la educación y de la psicología (P.V.P.: 2.800 pts.).

Los encontrará en los distribuidores de P.P.U.  
**PROMOCIONES Y PUBLICACIONES UNIVERSITARIAS**  
C/. Nicaragua, 100, 7.<sup>o</sup> 1.<sup>a</sup>, 08029-BARCELONA  
Teléf. (93) 239 91 37

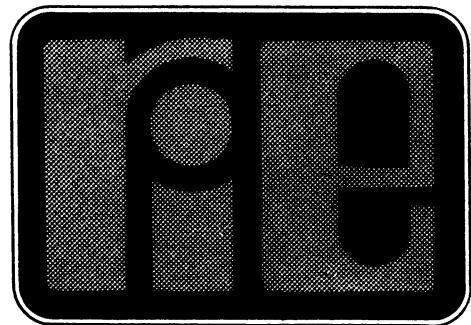


## III Seminario sobre modelos de investigación educativa

Con la colaboración de:

- INSTITUTO CIENCIAS DE LA EDUCACION-OVIEDO
- COMISION ASESORA DE INVESTIGACION CIENTIFICA Y TECNICA

**REVISTA  
INVESTIGACION  
EDUCATIVA**



## **GRUPOS DE COLABORACIÓN**

Departamento de Pedagogía Experimental, Terapéutica y Orientación  
**Universidad de Barcelona**

Departamento de Metodología y Tecnología Educativa  
**Universidad de Barcelona**

Departamento de Ciencias de la Educación  
Estudi General de Girona  
**Universidad Autónoma de Barcelona**

Instituto de Ciencias de la Educación (I.C.E.)  
**Universidad de Barcelona**

Departamento de Pedagogía.  
**Universidad de Granada**

Departamento de Pedagogía Experimental y Orientación  
**Universidad Complutense de Madrid**

Departamento de Pedagogía Experimental y Orientación  
**Universidad Nacional de Educación a Distancia. Madrid**

Departamento de Didáctica  
**Universidad Nacional de Educación a Distancia. Madrid**

Departamento de Pedagogía Experimental  
**Universidad de Murcia**

Departamento de Pedagogía Experimental  
**Universidad de Oviedo**

Instituto de Ciencias de la Educación (I.C.E.)  
**Universidad del País Vasco**

Departamento de Metodología Educativa  
**Universidad de Salamanca**

Departamento de Pedagogía  
**Universidad del País Vasco (San Sebastián)**

Departamento de Pedagogía Sistématica (Área de Pedagogía Experimental)  
**Universidad de Santiago de Compostela**

Departamento de Didáctica y Organización Escolar  
**Universidad de Sevilla**

Departamento de Pedagogía Experimental y Diferencial  
**Universidad de Valencia**

Vol. 3. N.º 6 2.º semestre 1985  
EDITA  
Asociación Interuniversitaria de  
Investigación Pedagógica  
Experimental (A.I.D.I.P.E.)

Director: Mario de Miguel  
Director ejecutivo:  
Benito Echeverría

**CONSEJO ASESOR:**  
Margarita Bartolomé  
Nuria Borrell  
Pilar Colas  
Iñaki Dendaluce  
Benito Echeverría  
José Fernández Huerta  
Fuensanta Hernández  
Mario de Miguel  
Arturo de la Orden  
Rosario de Pablo  
Ramón Pérez Juste  
José L. Rodríguez Diéguez  
Rafael Sanz Oro  
Francisco J. Tejedor  
Carmen Vidal

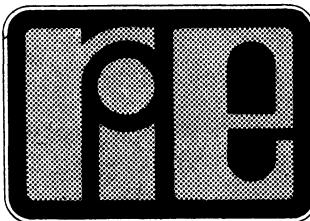
**CONSEJO DE REDACCIÓN:**

Margarita Bartolomé  
Rafael Bisquerra  
Inmaculada Bordas  
Flor Cabrera  
Julia V. Espin  
Jesús Garantó  
Juan Mateo  
Miguel Meler  
M.ª Luisa Rodríguez  
Mercedes Rodríguez  
Sebastián Rodríguez  
Delio del Rincón  
Pedro Sánchez  
Saturnino de la Torre

**ADMINISTRACIÓN,  
SUSCRIBTORES E  
INTERCAMBIO  
CIENTÍFICO**  
«Revista de Investigación  
Educativa».  
Dpto. P. Experimental.  
Terapéutica y Orientación  
Facultad de C. Educación  
Baldíri Reixach s/n. Bloque D. 3.<sup>º</sup>  
08028-BARCELONA  
Tels.: 240 92 00-08-09. Ext 212

**DISEÑO Y MAQUETACIÓN:**  
E. Molinero

**EDITOR:**  
Promoción y Publicaciones  
Universitarias  
Nicaragua, 100, 7.º 1.<sup>a</sup>  
08029-BARCELONA  
Teléf. 239 91 37



Vol. 3, N.º **6**

# REVISTA INVESTIGACION EDUCATIVA

## SUMARIO

**EDITORIAL** \_\_\_\_\_ 101

**EN PORTADA**

La investigación empírica sobre el rendimiento  
en España en la década 1975-1985 ·  
por M.ª Angeles Marín, Raquel Martínez,  
Nuria Rajadell \_\_\_\_\_ 103

**PONENCIAS**

Posibilidades de la metodología cualitativa US.  
Cuantitativa \_\_\_\_\_ 127  
por M.ª Teresa Anguera \_\_\_\_\_

Incidencia en la práctica educativa de la  
investigación en el rendimiento  
por Margarita Bartolomé \_\_\_\_\_ 145

Evaluación con respecto al criterio en matemáticas  
(Proyecto E.A.O. - TOAM)  
por Inmaculada Bordas \_\_\_\_\_ 155

Proyecto E.A.O. - TOAM. Estudio piloto sobre  
variables asociadas al rendimiento  
por Flor Cabrera, Juana Gómez \_\_\_\_\_ 171

Proyecto E.A.O. - TOAM  
Evaluación del rendimiento aritmético  
por Benito Echeverría \_\_\_\_\_ 183

Tendencias actuales de la documentación  
científica automatizada en la investigación educativa  
por Luis Lizasoain \_\_\_\_\_ 211

Meta-análisis sobre la predicción del rendimiento  
escolar  
por M.ª Paz Loscos \_\_\_\_\_ 221

Meta-análisis correlacional sobre estudios  
de rendimiento escolar en España  
por Juan Mateo \_\_\_\_\_ 236



Estrategias metodológicas en los estudios longitudinales por Mario de Miguel	252
Hacia una conceptualización del producto educativo por Arturo de la Orden	271
Modelos de investigación sobre el rendimiento académico. Problemática y Tendencias por Sebastián Rodríguez Espinar	284
Evaluación del rendimiento criterial versus normativa. Modelo de evaluación F.C.O. por Mercedes Rodríguez Lajo	304
Problemática de la enseñanza universitaria por Javier Tejedor	322

#### **COMUNICACIONES**

por Justo Arnal (339). Antonio R. Bartolomé (346).  
José Benito (352). Rafael Bisquerra (358).  
Leonor Buendía, Juan Ruiz (364). José Cajide y  
Xose Rubal (371). M.ª Díez Miralles y otros (386).  
Eduardo García Jiménez (390). José Luis  
García Llamas (395). Carmen Jiménez Fernández  
(401). A López Blasco, J. Marqués Ruiz,  
A. Martínez Sánchez (407). Juan de Pablos (412).  
J.T. Pastor (417). M.ª Dolores Peris (421).  
Antonio de Pro. Sebastián Sánchez (428).  
Carlos Rosales (435). Francisco de Salvador  
(440). Pedro Sánchez Algarra (445). Nicolás  
Seisdedos (449). Felipe Trillo (455).  
Luis Miguel Villar (462).

## Punto y seguido

Celebrar una reunión como la del «*III Seminario sobre Modelos de Investigación Educativa*» siempre es motivo de alegría y obligadas reflexión.

A todo ser humano, y especialmente a los que nos dedicamos a la investigación empírica en el campo educativo, nos alegra poder compartir los sueños, inquietudes y preocupaciones que animan nuestro quehacer. El tiempo es corto en estas reuniones y para alargarlo y hacerlo extensible a las personas con intereses similares, R.I.E. recoge en este número las ponencias y comunicaciones de este seminario, celebrado en Gijón, para analizar las «*Nuevas tendencias de la investigación sobre el rendimiento académico*».

Por primera vez estos Seminarios de Modelos, han dejado las plácidas aguas del mediterráneo catalán, que abrieron sus ojos, para introducirse con autonomía en las embravecidas olas del cantábrico.

Vieron la luz en 1981 (Barcelona) casi sin hacer ruido y con esa mezcla de debilidad y esperanza que acompaña a todo recién nacido.

Dieron sus primeros pasos en 1983 (Sitges), que sirvieron para descubrir el panorama investigador de nuestro «país educativo» y celebrar el nacimiento de la *Revista de Investigación Educativa*, altavoz de los análisis que allí se realizaron (Vol. II, N.º 2) y de trabajos posteriores.

Gijón, el año 1985 y los que vendrán hasta la celebración del IV Seminario en Santiago pueden ser históricos para una asociación como A.I.D.P.E., que pasa ahora por esos momentos necesarios de identidad, juventud y crisis.

Va a ser un bienio crucial, un tiempo crítico, que, tras la acumulación de los recursos de nacimiento, crecimiento y diferenciación, va a permitir con toda seguridad que el desarrollo de la investigación educativa en nuestro país asuma sus propias señas de identidad.

Hasta 1987 los equipos de investigación van a vivir esa tensión activa, esos ratos y compromisos adquiridos en Asturias, que entre todos vamos a potenciar con reuniones de trabajo reducidas, con canales de información ágiles y flexibles, pero sobre todo con horas y horas de trabajo reflexivo.



# En portada

Revista Investigación Educativa - Vol. 3 - N.º 6 - 1985 (P.103 - 126 )

## La investigación empírica sobre el rendimiento en España en la década 1975-1985.

por  
M.ª Angeles Marín  
Raquel Martínez  
Núria Rajadell

### PRESENTACIÓN

Uno de los objetivos de este III Seminario de Modelos es el de revisar las investigaciones empíricas realizadas en nuestro país en torno al tema del Rendimiento Educativo.

La presente comunicación pretende contribuir al logro de este objetivo, facilitando una síntesis de los estudios efectuados durante la última década referentes a esta temática.

La buena acogida que tuvo por parte de los asistentes al II Seminario de Modelos el resumen de los distintos trabajos agrupados por Líneas de Investigación nos ha animado a presentar una recopilación semejante que nos ayuda a reflexionar sobre el estado de la cuestión. Como es sabido por todos ya desde el I Seminario se planteó como objetivo prioritario la incidencia en la práctica educativa de la investigación del Rendimiento. Creemos que aparte de todos los trabajos que se presentarán en estos días, sería bueno tener en cuenta las investigaciones que aquí se citan en el momento en que hablaremos sobre este tema.

### INTRODUCCIÓN

Las fuentes que han servido de base para realizar el presente análisis son: las Tesis Doctorales y Tesinas de Licenciatura presentados en las distintas universidades españolas durante este década. Se consultaron las Fichas Resumen de la *Revista RIE* y se pidió colaboración a las diferentes Facultados; en algunos casos la recopilación se hizo directamente en la Universidad. Los datos correspondientes a las investigaciones proceden del estudio y revisión de 12 títulos de Revistas Educativas de nuestro país; de Ponencias o comunicaciones presentadas a Congresos; de los Planes de Investigación de la Red INCIE-ICE, y de otros

TABLA I  
Investigaciones según lugar de procedencia, año y carácter

UNIVERS	AÑO CARACT.			1975			1976			1977			1978			1979			1980			1981			1982			1983			
	D	L	O	D	L	O	D	L	O	D	L	O	D	L	O	D	L	O	D	L	O	D	L	O	D	L	O	D	L	O	
BARCELONA	4			2						1			1			1	2	2				4			6	1	5	1	3	2	3
BILBAO					1																										
EXTREMADURA										1																					
GRANADA																															
LA LAGUNA										1																					
MADRID				2	2	2		1					1			2	5	1		1	1	4	3	6	2	1	3	1	2	2	
MURCIA																							1								
NAVARRA																															
OVIEDO																															
SALAMANCA																															
SEVILLA																															
UNED																															
VALENCIA				1																											
VALLADOLID																															
ZARAGOZA																															
sin especificar																															

D = Tesis Doctoral (20)  
L = Tesis de Licenciatura (44)  
O = Otras Investigaciones (70)

planes de investigación de las diversas comunidades autónomas españolas. Se ha consultado también las fichas de líneas de investigación de la *Revista RIE*. Algunos de los trabajos analizados están pendientes de publicación.

Somos conscientes de que, a pesar del esfuerzo, es posible que no queden recogidas todas las investigaciones efectuadas en torno al tema que nos ocupa. De algunas investigaciones nos han faltado datos, bien relativos a la muestra, bien a las variables o al criterio empleado para medir el rendimiento. De aquellas que teníamos más datos se han contabilizado en los totales, pero no se han podido analizar en los núcleos temáticos. Otras no las hemos contabilizado por disponer sólo del autor y el título.

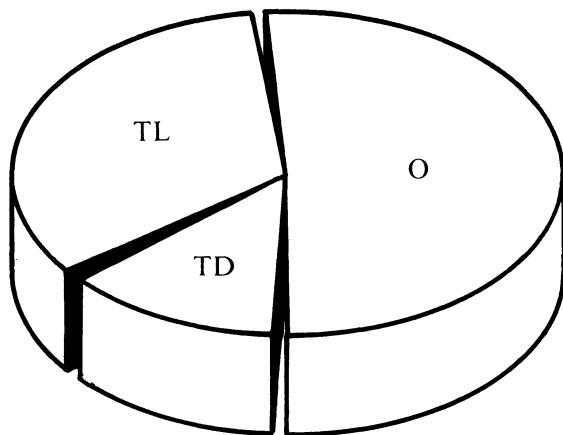
## DELIMITACIÓN DEL ESTUDIO

El total de trabajo analizados son 134. De ellos 22 están recopilados por el Dr. Benito Echeverría en su ponencia presentada al II Seminario de Modelos de Investigación. Aquí se presentan nuevamente distribuidos según los niveles temáticos que se han elaborado para este estudio.

La Tabla I muestra el reparto de las investigaciones según el lugar de procedencia, año y carácter.

En el Gráfico I se presentan el número de trabajos analizados por carácter.

GRAFICO I Número de investigaciones analizadas según su carácter

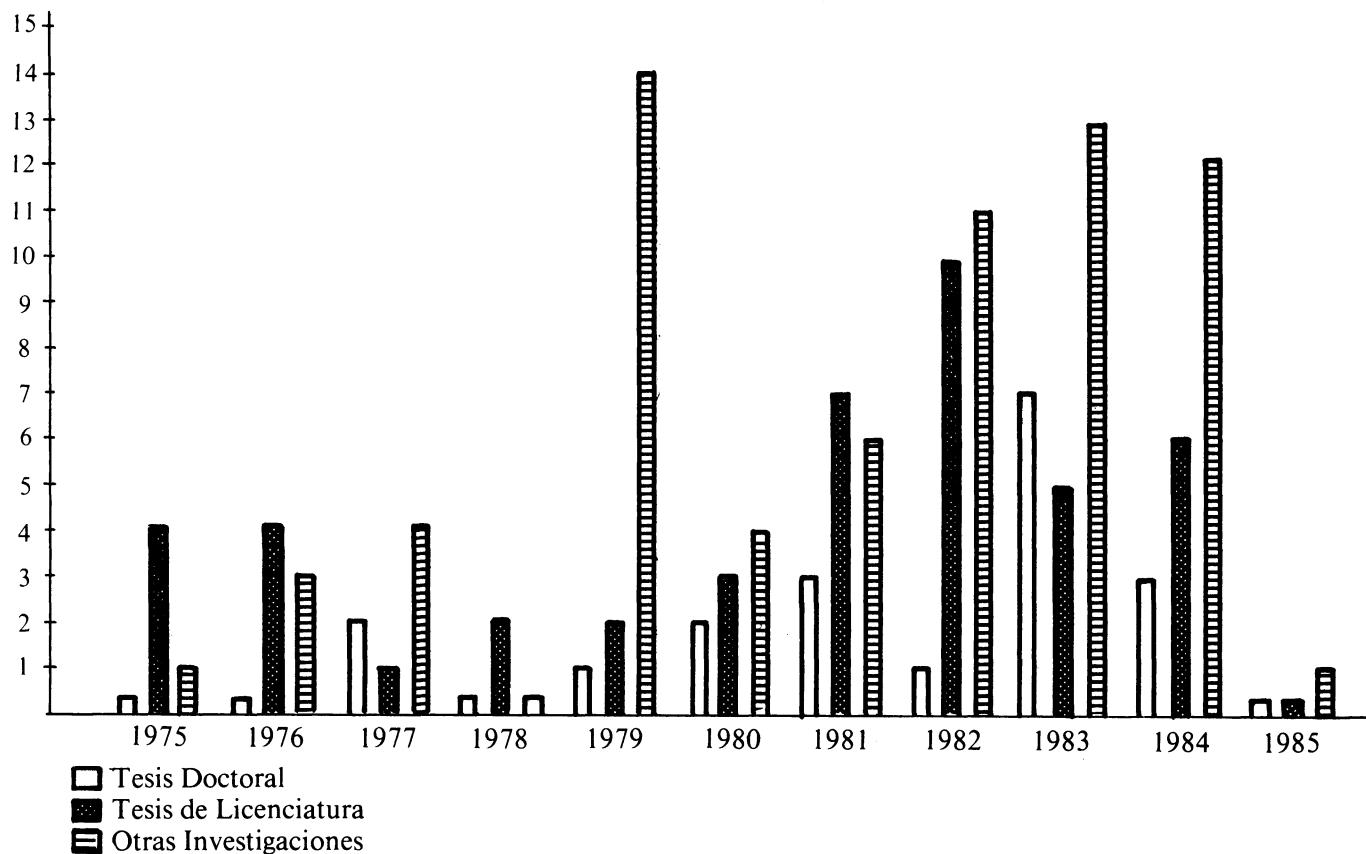


TD = Tesis Doctoral (14,18 %)

TL = Tesis de Licenciatura (32,83 %)

O = Otras Investigaciones (52,98 %)

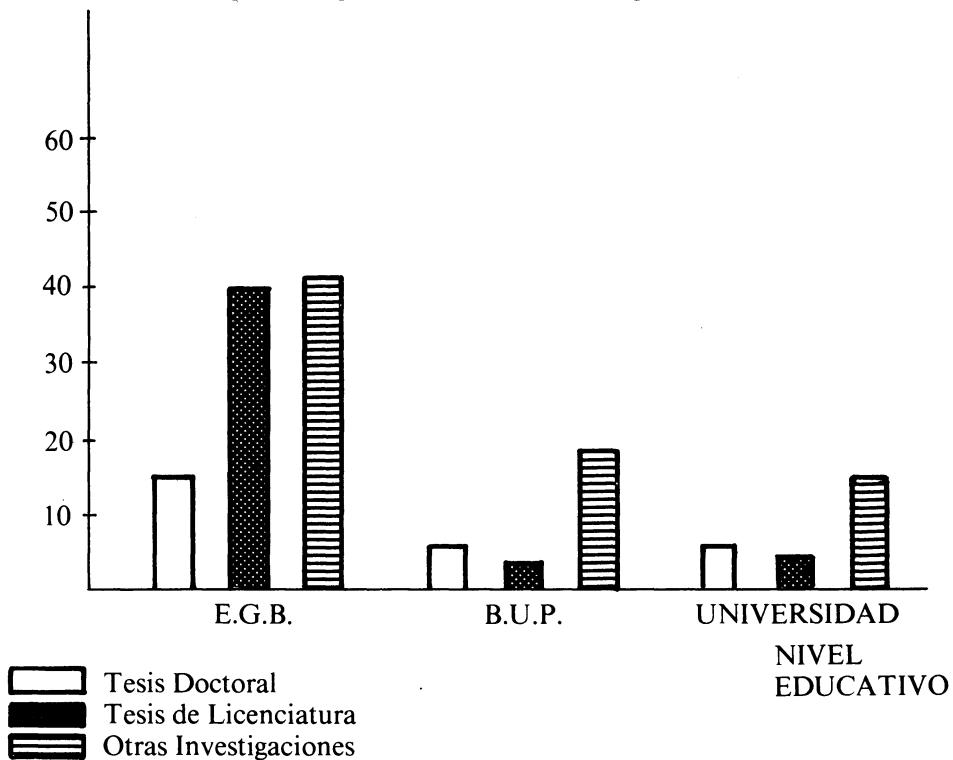
GRAFICO II  
Número de trabajos realizados por año, y glosados según su carácter



En el Gráfico II aparecen el número de trabajos hallados por año y desglosados según su carácter.

En el Gráfico III se recogen las investigaciones por niveles y carácter.

GRAFICO III  
Investigaciones por niveles educativos según su carácter



## CRITERIOS DE CLASIFICACIÓN

### 1. Clasificación temática

Para analizar los trabajos se ha seguido el esquema que figura en las fichas resumen. En el caso de las Tesis Doctorales y de Licenciatura anteriores a 1980, año en que se comenzaron a recopilar las fichas y líneas de investigación, se han elaborado unas guías semejantes que se guardan en el archivo de la *Revista de Investigación Educativa*.

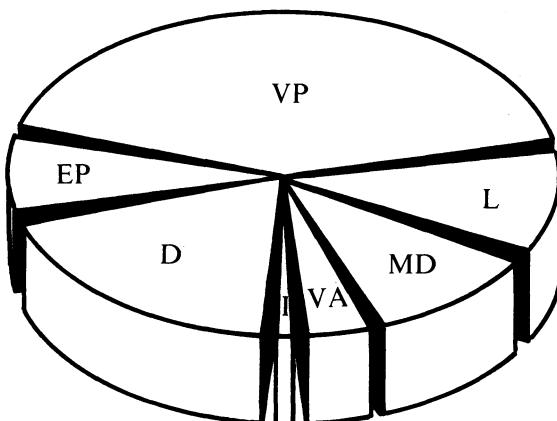
En un primer momento nos planteamos una clasificación temática en función de las variables empleadas en las investigaciones para establecer su relación con el rendimiento; separando por una parte aquellas que hacían referencia al alumno, tanto en lo que respecta a variables personales como socioeconómicas, y por otra las que se centraban más en la escuela.

Nos encontramos con una variedad de estudios tal que hacia difícil presentar ordenadamente los trabajos siguiendo este criterio, puesto que también se había de tener en cuenta el criterio con el que se mide el rendimiento.

Con el ánimo de proporcionar una información lo más detallada posible, hemos optado por una distribución temática que puede proporcionar una visión de la variedad de trabajos realizados diferenciados según el objetivo principal que se proponían. De esta forma se pueden apreciar aquellas áreas donde se ha centrado más el interés de las investigaciones y aquellas que han recabado menor atención. Aunque para la clasificación general temática se ha tenido en cuenta el objetivo principal de la investigación, dentro de cada núcleo se han considerado unos subnúcleos donde se contabilizan para el análisis aquellas investigaciones que tenían relación con este subnúcleo, aunque su objetivo principal perteneciera a otro.

En el Gráfico IV se presentan los porcentajes de cada núcleo temático.

GRAFICO IV  
Distribución porcentual de las investigaciones según su temática



D = Descriptiva general	(21,60 %)
EP = Elaboración de Pruebas	( 8,95 %)
VP = Variables Personales	(43,00 %)
L = Lenguaje	( 9,70 %)
MD = Métodos Didácticos	(11,19 %)
VA = Variables Ambientales	( 5,35 %)
I = Métodos de Investigación	( 2,23 %)

Ahora bien pensando que puede ser útil a la hora de la discusión tener un resumen más global siguiendo el criterio que se había propuesto en un principio, las conclusiones y prospectiva las plantearemos siguiendo este esquema.

## 2. Clasificación metodológica

Se ha seguido el guión adoptado en el Seminario anterior por el Dr. Benito Echeverría y como él apuntaba «únicamente con una finalidad práctica». (Véase *Revista Investigación Educativa*, Vol 2, N.<sup>o</sup> 1

La clasificación según el fin primordial que persigue ha sido:

- |                                 |                           |
|---------------------------------|---------------------------|
| a) Análisis y clasificación:    | - DESCRIPTIVAS            |
| b) Prueba de hipótesis:         | - EXPLICATIVAS            |
|                                 | - De casos                |
|                                 | - Selectivas              |
|                                 | - Correlacionales         |
|                                 | - Causales                |
|                                 | - PREDICTIVAS             |
|                                 | - EXPERIMENTALES          |
|                                 | - Cuasiexperimentales     |
| c) Intentos, logros, objetivos: | - INVESTIGACIONES ACTIVAS |
|                                 | - TOMA DE DECISIONES      |

## NUCLEOS TEMÁTICOS

### 1. Análisis general del Rendimiento (Descriptivo).

Los estudios de carácter descriptivo del Rendimiento como tal equivalen a un 21,6% y se distribuyen según la Tabla II.

**TABLA II** Distribución de las investigaciones por niveles y carácter \*

CARACTER NIVEL	TD	TL	O	TOTAL
EGB		7	10	17
ENS. MEDIAS	5		6	11
UNIVERSIDAD	1	1	6	8
<b>TOTAL</b>	<b>6</b>	<b>8</b>	<b>22</b>	<b>36</b>

\* Se consideran los niveles y no el número concreto de investigaciones, pues en este caso nos encontramos con tres de ellas de tipo longitudinal que abarcan más de un nivel.

Se pueden desglosar en cuatro subnúcleos temáticos:

A) El primero consistiría en la simple descripción de centros escolares (dos de EGB y uno universitario), analizando en ellos su organización, las áreas que se imparten e incluso algunas variables ambientales como aspectos que más pueden influir en el rendimiento de los alumnos.

Dos estudios concuerdan tanto en el nivel de EGB ya mencionado, como en ser Tesis de Licenciatura. Utilizan una metodología descriptiva empleando como indicadores del rendimiento únicamente las calificaciones escolares.

La otra investigación, centrada en una Facultad Universitaria, emplea una metodología de carácter predictivo. Es una Tesis Doctoral, y como indicadores no se contente simplemente con las calificaciones sino que además, aplica a los sujetos de estudio algunas pruebas de rendimiento en varias áreas escolares.

B) Otro apartado engloba el amplio tema del fracaso escolar, considerando también toda la faceta adaptativa del alumno a la escuela y a las materias formales que en ella se imparten, tema excesivamente trabajado en las revistas educativas de nuestro país a nivel meramente teórico, o con profusión de datos estadísticos poco analizados.

Este campo ha sido y continua siendo de sumo interés a lo largo de estos últimos años, sobre todo desde que entró en vigor la LGE en 1970. El número de trabajos mayor se encuentra en el nivel de EGB, sobre todo en el Ciclo Superior, aunque hay alguno que abarca hasta el COU.

En general se considera que se da fracaso escolar cuando existe un rendimiento bajo, medido a través de las calificaciones escolares. En los estudios analizados se intentan explicar las causas de este bajo rendimiento, bien a través del estudio de casos, (seleccionando sujetos con bajas calificaciones y viendo el peso de variables aptitudinales, de personalidad o sociales); bien mediante estudios correlacionales o predictivos que emplean este mismo tipo de variables como predictoras del rendimiento.

Un trabajo se centra exclusivamente en la relación entre la adaptación escolar medida a través de un cuestionario y el rendimiento en las diferentes materias, diferenciando niveles escolares (EGB y BUP) y el sexo. Otras investigaciones que encontramos son de carácter descriptivo del rendimiento en diversos zonas geográficas españolas. Por último hay una tesis de Licenciatura que se ocupa de la predicción del rendimiento en BUP y FP a partir de las calificaciones escolares de la EGB.

C) Otro subnúcleo sería el que trata toda la temática de la evaluación, y que en muchos aspectos se encuentra relacionado con el apartado anterior.

El tema de la evaluación, que ha despertado el interés en todos aquellos que se relacionan con el campo educativo, en nuestro estudio se refleja con un mayor porcentaje de investigaciones desde el año 1983, tanto en Tesis Doctorales, como en Tesis de Licenciatura como en artículos de carácter empírico. Las

diferentes investigaciones se encuentran en un porcentaje muy similar, aunque quizás donde menos ejemplos encontramos es en las Tesis de Doctorado.

El tema de la evaluación se ha trabajado en todos los niveles escolares, tanto en EGB, como BUP o COU, e incluso hallamos una artículo en el cuál se profundiza en toda la escolarización. FP sigue siendo el nivel escolar más olvidado por todos los investigadores.

Las metodologías utilizadas son diversas: hay investigaciones de carácter descriptivo sobre la influencia de diferentes métodos de evaluación, o sobre la disparidad de criterios evaluativos en el rendimiento. Un estudio cuasiexperimental para ver si la adecuación de criterios en cuanto a los contenidos por parte de los profesores tiene una repercusión en el rendimiento. Encontramos también unas investigaciones explicativas sobre modelos de evaluación del rendimiento en EGB, BUP y Universidad.

D) Y por último un trabajo, dentro de este núcleo de carácter general del rendimiento, en el que se estudian las causas del mayor rendimiento en las niñas que en los niños en los niveles de 5.<sup>º</sup> y 8.<sup>º</sup> de EGB. Se analizan variables aptitudinales, de hábitos de estudio, métodos de trabajo y personalidad.

## 2. Elaboración de pruebas

Dentro de este núcleo se han considerado todos aquellos trabajos que tenían en alguno de sus objetivos la construcción o adaptación de pruebas de rendimiento.

El nivel educativo con mayor objeto de atención ha sido la EGB.

Las investigaciones se distribuyen casi por un igual según su carácter. La Tabla III muestra el número de investigaciones por carácter y nivel.

Se observa una gran preocupación por la evaluación del rendimiento en lenguaje, y dentro de éste la lectura es el campo de atención preferente.

En este grupo de investigaciones hay una clara conexión temática y metodológica, siguiendo los pasos de la elaboración de instrumentos de medida. Son pruebas de diagnóstico del lenguaje (oral y escrito) o de rendimiento lector (comprensión, eficacia). En este último aspecto algunas son adaptaciones al catalán de tests existentes en castellano. Al introducirse la enseñanza del catalán, la necesidad de evaluar el proceso de aprendizaje y el rendimiento lector en niños bilingües ha motivado la realización de estas investigaciones. Estos trabajos contienen también estudios descriptivos del rendimiento lector en función de la clase social y de la lengua materna, del tipo de escuela, del sexo o de las aptitudes intelectuales.

Otros trabajos se han interesado por el rendimiento en áreas concretas como las Ciencias de la Naturaleza, las Matemáticas o la Física y Química. Son pruebas de diagnóstico o de rendimiento. En algunos casos se han elaborado para evaluar una metodología didáctica. Se presentan estudios descriptivos del rendi-

**TABLA III** Distribución de las investigaciones por niveles y carácter.\*

CARÁCTER NIVEL	TD	TL	O	TOTAL
EGB Inicial	1	1	1	3
EGB Medio	2	1	1	4
EGB Superior	1	4	1	6
ENS Medias			2	2
UNIVERSIDAD	1			1
<b>TOTAL</b>	<b>5</b>	<b>6</b>	<b>5</b>	<b>16</b>

\*Se consideran los niveles y no el número de investigaciones. En este caso es una sola investigación pero en los tres niveles de EGB.

miento en estas áreas en función de la clase social, la tipología escolar o el coeficiente intelectual.

### 3. Variables Personales

El 32% del total de las investigaciones analizadas está formando por ese amplio tema que detallamos por niveles y carácter en la Tabla IV.

**TABLA IV** Distribución de las investigaciones por niveles y carácter.

CARÁCTER NIVEL	TD	TL	O	TOTAL
EGB	7	7	15	31
ENS. MEDIAS	2	1	5	8
UNIVERSIDAD			4	4
<b>TOTAL</b>	<b>9</b>	<b>10</b>	<b>24</b>	<b>43</b>

Entre los trabajos analizados existe un buen número de estudios descriptivos o predictivos que utilizan como predictores variables personales (ya sean de tipo intelectual, aptitudinal, de personalidad, motivación, adaptación). Constituyen investigaciones exploratorias que se proponen observar la variabilidad del rendimiento (criterio) medido por las calificaciones escolares a partir de la variabilidad de los predictores, medidos por medio de test de inteligencia general aptitudes y personalidad.

En este apartado se recogen sólo aquellas investigaciones que utilizan los predictores citados como variable independiente; pero hay otro gran número de estudios en los que aparecen como variables intervintivas.

La distribución es bastante semejante para las Tesis de Doctorado y de Licenciatura, aunque el mayor número de trabajos lo constituyen investigaciones que no entran en estas categorías. El nivel más estudiado es el de la EGB.

Aunque el modelo seguido por la mayoría de los trabajos es el expuesto anteriormente, hay algunas investigaciones que se salen un poco de esta norma, bien por los instrumentos de medida utilizados con los predictores, bien en el planteamiento. Las comentamos brevemente.

Una Tesis de Licenciatura que se plantea el diagnóstico de los factores cognitivos por medio de un modelo operatorio y emplea como medida del rendimiento en matemáticas las calificaciones y una evaluación cualitativa.

Una Tesis Doctoral que introduce junto a las variables aptitudinales, la comprensión lectora como predictor, empleando como criterio pruebas de rendimiento en matemáticas elaboradas por la autora.

Otras plantean la influencia de los estilos cognitivos del alumno y profesor en el rendimiento.

Por último una investigación que analiza la influencia de métodos de enseñanza diversos en el rendimiento de los sujetos dependientes e independientes de campo. Introduce un nuevo tipo de medida, como la anterior, para la dimensión cognitiva. El criterio está medido a través de pruebas de recuerdo, inducción y síntesis.

El grupo de estudios que se centran más en las variables de personalidad como predictores difieren en cuanto a los objetivos y metodología. Existen unas de tipo descriptivo que estudian la afectividad desde la dimensión relacional, empleando como criterio las calificaciones.

Otros estudios predictivos que emplean variables de personalidad medidas a través de test proyectivos, autoevaluaciones, evaluación docente y aptitudes como predictores y las calificaciones finales, globales o por áreas como criterios del rendimiento.

Hay un estudio quasiexperimental para explicar la incidencia de los factores de personalidad en el rendimiento escolar. La novedad que presenta consiste en que se aparta de los modelos correlacionales introduciendo el análisis de perfiles para las dimensiones de personalidad. Esta técnica se ha empleado en estudios posteriores. Durante este Seminario tendremos ocasión de profundizar más en ella, viendo las ventajas que ofrece sobre los modelos correlacionales tradicionales, cuando tratemos el tema de los modelos de análisis e investigación del rendimiento en EGB y EEMM.

#### 4. Lenguaje

Un 9,7% de los trabajos analizados corresponden a esta área. Si estos se añaden todos aquellos que tienen como objetivo principal la elaboración de pruebas

para evaluar el rendimiento lector, es uno de los temas que agrupa mayor número de investigaciones.

Los datos de la Tabla V indican la distribución por niveles y carácter.

**TABLA V** Distribución de las investigaciones por niveles y carácter\*

CARÁCTER NIVEL	TD	TL	O	TOTAL
EGB. Inicial	2	7	2	11
EGB. Medio	3	7	1	11
EGB. Superior	2	8	1	11
UNIVERSIDAD	1			1
<b>TOTAL</b>	<b>8</b>	<b>22</b>	<b>4</b>	<b>34</b>

\* Dado que hay cierto número de investigaciones realizadas en fecha anterior a la actual división de EGB, la distribución de la muestra es, en algunos casos, correspondiente a los dos niveles de la división actual. Se ha optado por contabilizar estas investigaciones en ambos niveles.

El tema de la lectura ocupa una atención preferente, bien analizando las dificultades e intentando explicar los factores que las originan, ya sea con estudios predictivos, o viendo la eficacia de diferentes métodos de enseñanza. Aparecen también algunos trabajos que analizan las dificultades del lenguaje escrito.

Los estudios diferenciales o compensatorios que analizan las dificultades lectoras son investigaciones descriptivas en las que se tienen en cuenta variables personales (ambientales o psicológicas) o educativas.

Hay dos estudios predictivos del rendimiento lector en función de los aspectos madurativos medidos a través de pruebas de maduración lectora, emplean la ecuación de regresión.

Hay otra serie de trabajos que intentan mejorar el rendimiento lector a través de la estimulación de algún factor asociado con el aprendizaje de la lectura; y dos investigaciones que tratan de probar la eficacia de un método concreto de enseñanza de la lectura, uno de ellos realizado con niños de educación especial. La metodología empleada es de carácter diverso: explicativa, estudio de casos y cuasiexperimental.

## 5. Métodos-Medios Didácticos

Hemos englobado dentro de este núcleo temático los trabajos en los que se intenta probar la eficacia de diferentes métodos o medios didácticos para mejorar el rendimiento de los alumnos.

Se incluyen en el conjunto total los que se han comentado anteriormente referentes al lenguaje.

La distribución de estas investigaciones se presenta en la tabla VI.

TABLA VI Distribución de las investigaciones por niveles y carácter.

CARACTER NIVEL	TD	TL	O	TOTAL
EGB	1	5	5	11
ENS. MEDIAS			1	1
UNIVERSIDAD			3	3
<b>TOTAL</b>	<b>1</b>	<b>5</b>	<b>9</b>	<b>15</b>

Predomina el nivel de la EGB, y son artículos, investigaciones y tesis de licenciatura.

No analizamos los trabajos relacionados con el tema del lenguaje por haberse tratado de ellos en otro apartado. De los restantes mencionamos alguno.

Una investigación cuasiexperimental sobre la eficacia del ordenador, para ver en qué medida este tipo de enseñanza mejora el rendimiento de los alumnos. En el presente Seminario tendremos ocasión de profundizar este tema cuando estudiemos el programa TOAM.

La medida en que la enseñanza programada mejora la calidad de la enseñanza.

Y estudios sobre la validez de dos modelos tecnológicos para mejorar el rendimiento en el aprendizaje del inglés y la geografía.

Por último un grupo de trabajos que estudian la influencia de agrupaciones de alumnos, de las interacciones profesor-alumno, o diferentes sistemas de organización dentro de la clase, sobre el rendimiento.

## 6. Variables ambientales

En este apartado hay 19 trabajos que se proponen investigar la influencia de diferentes variables ambientales (familia, clase social, utilización del transporte, información de los padres, asistencia a clase, instrucción de los padres...) sobre el rendimiento.

Son en su mayoría estudios de tipo descriptivo o explicativo. Están realizadas en el nivel de EGB, aunque hay alguna en la Universidad.

## 7. Métodos de investigación

Y por último consignamos en este apartado trabajos que están relacionados con el tema por su propósito final. Algunos podrían estar incluidos en el tema de la elaboración de pruebas, pero no están específicamente centrados en la construcción sino que sus objetivos consisten o bien en la revisión de técnicas para el análisis de ítems en las pruebas de evaluación formativa planteando el problema del análisis de ítems en las pruebas criteriales, utilizando el análisis factorial, o bien en la construcción de tablas de predicción.

La siguiente presenta una serie de precisiones con relación a las metodologías utilizadas con más frecuencia para analizar el rendimiento, señalando, a partir de los datos de investigaciones concretas, sus ventajas y limitaciones, proponiendo modelos alternativos.

Tendremos ocasión de trabajarlos en más profundidad en las sesiones previstas para ello durante el Seminario.

## CONCLUSIONES. PROSPECTIVA

El análisis de las investigaciones realizadas sobre el rendimiento en esta década en las universidades españolas nos ha dado una visión general en cuanto a los aspectos más tratados y los modelos y metodologías empleados con más frecuencia.

Antes de finalizar esta exposición queríamos presentar un breve resumen de las principales tendencias que se apuntan en ellas, analizar los vacíos y aquellos aspectos más trabajados.

El mayor volumen de trabajos se registra a partir de 1981, y el nivel que ha recabado mayor atención por parte de los investigadores es la EGB.

A partir de la puesta en marcha de la Ley General de Educación en 1970 se despertó un creciente interés por el análisis del rendimiento escolar. Son numerosas las publicaciones realizadas a efectos divulgativos; el elevado número de trabajos de investigación que se han analizado confirma este interés.

Una observación detallada de las variables utilizadas como predictores en las diferentes investigaciones señala que la inteligencia, aptitudes y personalidad son los aspectos en los que hace más incidencia.

Es menor el número de investigaciones en las que se emplean como predictores variables ambientales, del medio, escolares, familiares,...

Si consideramos la mención del criterio esta está efectuada a través de las calificaciones en su mayor parte. Son pocos los trabajos destinados a elaborar pruebas de rendimiento. El área que más ha profundizado este aspecto es la del lenguaje, en el tema de la lectura más concretamente. Contamos con pruebas para el diagnóstico del estado inicial del aprendizaje lector, para la medición de la comprensión lectora, del lenguaje oral, del vocabulario, la lectura oral y la ortografía en los niños que finalizan el Ciclo Inicial (en castellano y catalán); de eficiencia lectora para el Ciclo Superior y Universidad.

Se observa un vacío a nivel de pruebas para medir la expresión oral y escrita. Es un campo en el que está trabajando y del que existen algunas pruebas que están siendo objeto de investigación. En el presente seminario hay alguna comunicación que abordará el tema de la expresión oral.

Los métodos correlacionales y la ecuación de regresión son los empleados con mayor frecuencia en estudios de tipo descriptivo, explicativo o predictivo. Existen también algunos trabajos cuasiexperimentales.

### RELACION DE INVESTIGACIONES RECOPILADAS

#### TESIS DOCTORALES

- ANDRES MIRALLES, V. (1983): *Comprensión del lenguaje escrito para el éxito en matemáticas. Estudio experimental casi-didáctico realizado en el ciclo medio en castellano y catalán*. Universidad de Barcelona. Barcelona.
- BAQUERO REY, E. (1977): *Estudio experimental de las variables que influyen en el rendimiento escolar*. Universidad Complutense. Madrid.
- BISQUERRA ALZINA, R. (1980): *Desarrollo didáctico de la eficiencia lectora*. Universidad de Barcelona. Barcelona.
- BORDAS ALSINA, I. (1979): *Diagnóstico matemática en EGB*. Universidad de Barcelona. Barcelona.
- CABRERA RODRIGUEZ, F. (1983): *Elaboración de una batería de diagnóstico analítico de lectura*. Universidad de Barcelona. Barcelona.
- COLAS BRAVO, P. (1984): *Análisis experimental de métodos y técnicas didácticas interaccionados con algunas características de los alumnos*. Universidad de Sevilla.
- COMES NOLLA, G. (1984): *Evaluación de la lectura comprensiva con sistema taxonómico*. Universidad de Barcelona.
- DEIROS GARCIA, G. (1983): *Comprensión lectora en el ciclo medio o final de la primera etapa de EGB. Estudio experimental en catalán y castellano*. Universidad de Barcelona.
- FREIXA BLANXART, M. (1983): *El condicionamiento social de las aptitudes intelectuales y su influencia sobre el rendimiento escolar*. Universidad de Barcelona.
- GARCIA LLAMAS, J. L. (1985): *Evaluación del rendimiento académico en la UNED*. Universidad Nacional de Educación a Distancia. Madrid.
- LEBRERO BAENA, M.ª. P. (1983): *Influencia de un método de lectura y escritura en el rendimiento de estas áreas*. Universidad Complutense. Madrid.

- LUIS RODRIGUES, E. (1983): *Estilo cognitivo, actitud del profesor hacia los alumnos individualmente considerados y rendimiento escolar*. Universidad Complutense. Madrid.
- MARTINEZ ARIAS, M. R. (1977): *Aptitudes, personalidad y rendimiento en matemáticas*. Universidad Complutense. Madrid.
- MATEO ANDRES, J. (1980): *Los rasgos de la personalidad, su incidencia en el rendimiento escolar*. Universidad de Barcelona.
- PRIETO SANCHEZ, D. (1981): *Dislexia y desarrollo del esquema corporal*. Universidad de Murcia.
- RAMOS DA SILVA CARVALHO, M.L. (1981): *Tablas de predicción para segunda etapa de E.G.B.* Universidad Complutense. Madrid.
- RUBIO, M.T. (1984): *Desarrollo de un modelo tecnológico para optimizar el aprendizaje del inglés en 7.º de EGB*. Universidad de Barcelona.
- TOURON FIGUEROA, F.J. (1982): *Aportación al conocimiento de los alumnos que acuden a la carrera de ciencias biológicas en la Universidad de Navarra*. Universidad de Navarra.
- VIDAL XIFRE, M.ª C. (1983): *La evaluación en el BUP. Problemas ocasionados por la diversidad de criterios*. Universidad de Barcelona.

#### TESIS DE LICENCIATURA

- AMAT CODINA, D. (1975). *La educación psicomotriz y los procesos de madurez lectora y escritora*. Universidad de Barcelona.
- ARTAL SERRAT, M. - FLAQUE CENTELLES, R. (1980): *Test de comprensió lectora en català*. Universidad de Barcelona.
- CAMINA DURANTEZ, M. A. (1977): *Nivel de información y rendimiento escolar en tercero de E.G.B.* Universidad Complutense. Madrid.
- CAMPS, T. - EGEA, M. A. (1982): *La reeducación de los trastornos de percepción visual para el aprendizaje lector*. Universidad de Barcelona.
- CARDONA HERNANDEZ, M. A. (1982): *La eficiencia lectora en los alumnos de segunda etapa de EGB de Verdún-Roquetas*. Universidad de Barcelona.
- CARRERA, C. y otros (1981): *Análisis de la predicción lectora*. Universidad de Barcelona.
- CASANOVAS DANES, S. - MADURELL RAMIS, J. (1981): *Metodología lúdico-lectora en educación especial*. Universidad de Barcelona.
- CERDA MARTIN, M. C. (1984): *Ánalisis diferencial del rendimiento escolar en EGB en función de la variable sexo*. Universidad de Valencia.
- CIRIA BENEDE, J. - GONZALEZ SOLER, R. (1980): *El fracaso escolar en el área del lenguaje escrito. Investigación operativa*. Universidad de Barcelona.

- CLOPAS FISA, A. - PELIQUIN PEREZ, F. (1982): *Análisis de las dificultades lectoras en 2.º de EGB y provisión de una metodología correctiva*. Universidad de Barcelona.
- COMES NOLLA, G. (1978): *Aportación al bilingüismo por medio de test de lectura comprensiva*. Universidad de Barcelona.
- COROMINAS ROVIRA, E. (1981): *Predictió del rendiment en alumnes de 8è. d'EGB*. Estudi longitudinal. Universidad de Barcelona.
- DE LA ROSA RAMIREZ, M. - DIEZ PRIETO, M. P. (1975): *Diagnóstico y tratamiento de las dificultades lectoras en niñas de 4.º, 5.º y 6.º de EGB*. Universidad de Barcelona.
- DE MATEO PEREZ, J. M. - GONZALEZ BENITO, B. (1975): *Estudio diagnóstico de escolares que fracasan*. Universidad de Barcelona.
- ESTANY BASSA, S. (1981): *Bateria predictiva i diferencial per a F.P.-I*. Universidad de Barcelona.
- FERNANDEZ PEREZ, M. D. (1978): *Relación entre emotividad y cálculo a nivel de 5.º de E.G.B.* Universidad Complutense. Madrid.
- FERNANDEZ VARA, M. V. (1982): *La influencia de la familia con anormalidades en el rendimiento escolar*. Universidad Complutense. Madrid.
- GALLARDO ALONSO, M. D. (1976): *Afición lectora y su relación con el rendimiento escolar*. Universidad Complutense. Madrid.
- GARCIA BELTRAN, M. T. (1981): *Estudio descriptivo del rendimiento académico en un centro educativo de EGB*.
- GARCIA JIMENEZ, E. (1983): *Evaluación diagnóstica para el diseño instruccional en el área de ciencias de la naturaleza en EGB*. Universidad de Sevilla.
- GARCIA LOPEZ, J. (1982): *La motivación escolar en un modelo predictivo del rendimiento*. Universidad de Barcelona.
- GASPAR ANGUITA, P. (1981): *Influencia de la retroalimentación en el rendimiento académico*. Universidad Complutense. Madrid.
- GESTOSO ACOSTA, J. A. (1981): *La segunda etapa de EGB en la provincia de Valencia*. Universidad de Valencia.
- GOMEZ MOLINA, J. R. (1981): *Estudio descriptivo y predictivo de las evaluaciones de BUP y CÓU en el País Valenciano*. Universidad de Valencia.
- GOMEZ MUÑOZ, R. (1980): *Cuantificación del vocabulario en quinto y octavo de EGB, de los barrios periféricos de Córdoba y su incidencia en el rendimiento escolar*. Universidad Complutense. Madrid.
- GONZALEZ GARCIA, P. (1976): *Personalidad, motivación, inteligencia y rendimiento académico*. Universidad Complutense. Madrid.
- GUEVARA MARTINEZ, A. (1983): *Determinación de los factores madurativos condicionantes del aprendizaje de la lectura*. UNED. Madrid.

- GUTIERREZ GOMEZ, M. (1976): *La televisión y los niños: su relación con el rendimiento escolar.* Universidad Complutense. Madrid.
- HERNANDO MARTINEZ, M. R. - RODRIGUEZ QUIÑONES, A. (1982): *Escuela-Taller. Una experiencia psicopedagógica en el tratamiento de la inadaptación escolar.* Universidad de Barcelona.
- HERRERO RIVAS, D. N. (1982): *Implicación de la afectividad en el rendimiento educativo.* Universidad de Barcelona.
- JIMENA MORO, M. (1984): *El ordenador personal y la educación: aplicaciones.* UNED. Madrid.
- JIMENEZ GONZÁLEZ, F. (1983): *Construcción y validación de una prueba de objetivos de desarrollo en el área de ciencias de la naturaleza.* UNED. Madrid.
- MAÑA ALVARENGA, E. - SANCHEZ POVEDANO, N. (1982): *Exploración de algunos factores psicológicos cognitivos subyacentes al aprendizaje de las matemáticas.* Universidad de Barcelona.
- MARQUES SIMON, P. (1984): *Una aproximación diferencial al proceso de fracaso escolar y sus causas.* Universidad de Valencia.
- MARTIN NUÑEZ, C. (1983): *La escritura en segunda etapa de EGB: dificultades y recuperación.* Universidad de Valencia.
- MEDRANO ANDRADE, A. (1984): *La función didáctica de la evaluación.* Universidad Complutense. Madrid.
- MEDRANO CALLEJAS, G. (1984): *Algunos determinantes del rendimiento académico contrastados en alumnos transportados y no transportados en dos zonas de Cuenca.* Universidad de Murcia.
- PALOMAR MARTINEZ, I. (1978): *Relación entre algunas aptitudes mentales y algunos tipos de rendimiento.* Universidad Complutense. Madrid.
- PALLARES SOLER, M. (1976): *Diagnóstico y tratamiento de las dificultades lectoras en niñas bilingües de 4.º, 5.º y 6.º de EGB.* Universidad de Barcelona.
- PARRILLA LARIO, T. J. - VIÑETA SOLER J. (1976): *Dimensiones mentales básicas en la evaluación de 5.º de EGB.* Universidad de Barcelona.
- PARRA MARTINEZ, J. (1982): *Factores ambientales que influyen en el rendimiento escolar desde la perspectiva de los profesores de EGB de la región de Murcia.* Facultad de Filosofía y Ciencias de la Educación. Murcia.
- PLAZA GIL, M.ª I. (1982): *Estudio en torno a la predicción del rendimiento escolar.* Universidad Complutense. Madrid.
- PONS CLAR, P. (1975): *Eficacia del método global-analítico para la enseñanza de la lectura y escritura en catalán.* Universidad de Barcelona.
- RODRIGUES DIAS, E. L. (1982): *Estilos cognitivos, características personales y rendimiento escolar de los alumnos.* Universidad Complutense. Madrid.

- SAHUQUILLO DIAZ, D. (1979): *Relación entre la posesión de conceptos básicos y el rendimiento en primero de EGB*. Universidad Complutense de Madrid.
- SAURA CALIXTO, P. (1984): *Influencia de los medios tanto materiales como de estímulo y apoyo en el rendimiento escolar*. Universidad de Murcia.
- VAL ALONSO, E. (1982): *Incidencia de la familia en el fracaso escolar*. Universidad Complutense. Madrid.
- YOLDI GARCIA, S. (1984): *Clase social y lenguaje diferencial en la morfología del lenguaje escrito en niños pertenecientes a distinta clase social*. UNED. Madrid.

#### OTRAS INVESTIGACIONES

- AVIA, M. D., MORALES, F. y RODA, R. (1982): "Determinantes del rendimiento académico". Ciudadado por Carabaña, J.: *Temas de Investigación Educativa*. 2. MEC. 1982.
- BARTOLOMÉ PINA, M. (1983): "Análisis descriptivos y diferenciales del lenguaje en la EGB". *Revista de Investigación Educativa*. N.º 2. pp: 205-239.
- BUCETA FERNANDEZ, J. M., GARCIA-ALCAÑIZ, E. y PARRON SOLLEIRO, P. (1982): "Influencia de la situación familiar de los padres en el rendimiento escolar y la inteligencia de sus hijos: Estudio experimental con niños de ocho años". *Revista de Psicología General y Aplicada*. Vol. 37 (3), pp: 549-556.
- CABALLERO CABALLERO, A. (1979): "Estructura de las variables psicopedagógicas de los escolares en el nivel tercero de EGB", en "Dimensiones psicopedagógicas y predicción del rendimiento escolar". *Revista de Psicología General y Aplicada*. Vol. 34. N.º 160-161. pp: 907-927.
- CABERO PEREZ, M. V. y MORO BERIHUETE, M. P. (1984): *Ecuaciones de predicción de lecturabilidad en lengua castellana en los niveles de EGB, BUP, FP y Enseñanza Universitaria*. Santiago de Compostela: VIII Congreso Nacional de Pedagogía.
- CABRERA RODRIGUEZ, F. (1984): "Estudio inicial de la lectura. Evaluación diagnóstico-analítica del rendimiento lector". *Revista de Investigación Educativa*. N.º 3. pp: 19-39.
- CARABAÑA, J. (1979): "Origen social, inteligencia y rendimiento académico al final de la EGB". *Temas de Investigación Educativa*. MEC. Madrid. pp: 29-71.
- CID, R., BERNARD, J. A., ESCUDERO, T. y VALDIVIA, C. (1977): *Rasgos de madurez y éxito en las pruebas de acceso a la Universidad*. V Plan de Investigación. ICE. Univers. de Zaragoza.
- CODINA BAS, J. B. (1983): "Influjo de algunas variables sociológicas en el éxito-fracaso académico". *Bordón*. Septiembre-Octubre. N.º 0. pp: 28-30.
- CRESPO VASCO, J. (1979): "Dimensiones psicopedagógicas y rendimiento académico a nivel de 5.º curso de EGB", en "Dimensiones psicopedagógicas y predicción del rendimiento escolar". *Revista de Psicología General y Aplicada*. Vol. 34. N.º 160-161.
- CHORRO, J. L. (1982): "Enseñanza programada y calidad del rendimiento". *Revista de Investigación Educativa*. N.º 2. pp: 307.

- DE LA ORDEN HOZ, A. (1983): "La investigación sobre la evaluación educativa". *Revista de Investigación Educativa*. N.º 2. pp: 240-248.
- DPTO. PEDAGOGIA EXPERIMENTAL Y ORIENTACION (Universidad Complutense de Madrid) (1983): "Estilo cognitivo, elección de carrera y rendimiento de los estudiantes universitarios". *Revista de Investigación Educativa*. N.º 0. pp: 28-30.
- DOMINGO FRISA, M. A. y otros. (1983): "Análisis del fracaso escolar en relación a la procedencia social de los alumnos". *Revista de Sociología: Revista de Investigaciones Sociológicas de la Asociación Castellana de Sociología*. 2.
- ESCUDERO ESCORZA, T. (1981): *Selectividad y rendimiento académico de los universitarios. Condiciones psicológicas, sociológicas y educacionales*. Zaragoza. ICE.
- ESPINOSA, M. J. (1979): "El problema de las dimensiones de la inteligencia en primero de EGB: Aproximación factorial", en "Dimensiones psicopedagógicas y predicción del rendimiento escolar". *Revista de Psicología General y Aplicada*. Vol. 34. N.º 160-161. pp: 907-927.
- FDEZ POZAR y otros. (1980): *Agrupamientos flexibles de alumnos y profesores: Un nuevo concepto de organización de centros de EGB*. I Plan de Investigación. ICE. Univ. de Sevilla.
- GARCES CAMPOS, R., ESCUDERO ESCORZA y otros. (1985): *Los alumnos de BUP-COU y FP en Zaragoza: Características socioeducativas*. Delegación de Enseñanza y Escuelas Infantiles. Ayuntamiento de Zaragoza. Zaragoza.
- GARCIA-ALCAÑIZ, E. y BUCETA, J. M. (1981): "Influencia de la institucionalización durante los primeros años de vida en el rendimiento escolar y la inteligencia". *Análisis y Modificación de Conducta*. Vol. 7. N.º 15.
- GARCIA HOZ, V. (1979): "Tablas de Predicción del rendimiento escolar: concepto, construcción y uso". *Revista Española de Pedagogía*. N.º 146, pp: 3-18.
- GARCIA YAGUE, J. (1979): "Resultados de nuestros análisis factoriales sobre la estructura factorial de la inteligencia en tercero de EGB", en "Dimensiones psicopedagógicas y predicción del rendimiento escolar". *Revista de Psicología General y Aplicada*. Vol. 34. N.º 160-161, pp: 907-927.
- GARRETA, N. (1982): "Diferencias de autoestima en inteligencia y trabajo, en alumnos de quinto y octavo de EGB". *Temas de Investigación Educativa*, 2. MEC. Madrid.
- GIL CRIADO, A. (1979): "Algunos resultados de nuestras experiencias con los diferentes modelos de análisis factorial en el estudio de las dimensiones psicopedagógicas", en "Dimensiones Psicopedagógicas y predicción del rendimiento escolar". *Revista de Psicología General y Aplicada*. Vol. 34. N.º 160-161, pp: 907-927.
- GOMEZ MOLINA, J. R. (1982): "Estudio descriptivo de las evaluaciones de BUP y COU en el País Valenciano". *Bordón*. Enero-Febrero N.º 241, pp: 19-44.
- GONZALEZ GARCIA, E. M. (1981): "Un método de análisis del rendimiento académico en la asignatura de Ciencias Naturales del curso primero de Bachillerato". *Bordón*. Enero-Febrero. N.º 236, pp: 77-90.
- HERNANDEZ VILLA, S. (1981): "Adaptación familiar del adolescente y su influencia en el rendimiento escolar", en PEREZ SERRANO, G. *Origen social y rendimiento escolar*. Centro de Investigaciones Sociológicas. Madrid. 1981.

- INCIE (1976): *Determinantes del rendimiento académico*. MEC. Madrid.
- INSPECCION EGB de ZARAGOZA (1984): “Análisis del fracaso escolar en relación a las áreas curriculares”, en MOLINA GARCIA, S. y GARCIA PASCUAL, E.: *El éxito y el fracaso escolar en la EGB*. Laia. Barcelona.
- JIMENEZ FERNANDEZ, C. (1983): “Evaluación del rendimiento en la Universidad Nacional de Educación a Distancia”. *Revista Española de Pedagogía*. Año XLI. N.º 162. Octubre-Diciembre.
- JIMENEZ FERNANDEZ, C. (1984): “Acotaciones al cuestionario de adaptación escolar”. *Revista de Investigación Educativa*. N.º 4. pp: 133-144.
- LAFITE, M., LOPEZ, J.I. y PUJOL, A. (1983): “Geografía humana: búsqueda de soportes audiovisuales (la ciudad)”. *Revista de Investigación Educativa*. N.º 4. pp: 185.
- LAZARO MARTINEZ, A. J. (1979): “Incidencia escolar de la estimación social de los alumnos”, en Dimensiones psicopedagógicas y predicción del rendimiento escolar”. *Revista de Psicología General y Aplicada*. Vol. 34. N.º 160-161. pp: 907-927.
- LERA ALDONZA, O. “Personalidad, motivación y rendimiento académico”, en PEREZ SERRANO, G.: *Origen social y rendimiento escolar*. Centro de Investigaciones Sociológicas. Madrid.
- LERENA, C. (1975): *Escuela, ideología y clases sociales en España*. Ariel. Barcelona.
- LOPEZ, E. (1982): “Instrucción de los padres y desarrollo intelectual de sus hijos”. *Bordón*. Enero-Febrero. N.º 241. pp: 67-87.
- LOPEZ LOPEZ, E. (1982): “Asistencia a clase y rendimiento”. *Revista Española de Pedagogía*. Año XL. N.º 155. Enero-Marzo.
- LOPEZ MARTINEZ, J. (1984): “Investigación sobre el fracaso escolar en la Formación Profesional de los centros públicos de la Comunidad Autónoma de Madrid”. *Revista de Educación*. N.º 274. pp: 93-114.
- LORCA, Grupo de Investig. Educ. (1983): “Diseño de evaluación de un seminario de física y química a nivel local”. *Revista de Investigación Educativa*. N.º 3. pp: 93.
- MARTINEZ, A. (1979): “Estudio experimental del rendimiento académico en distintas situaciones de aprendizaje”. *Revista Española de Pedagogía*. N.º 144. pp: 3-32.
- MARTINEZ SANCHEZ, A. (1980): “Estudio analítico del rendimiento académico de los grupos”. *Revista Española de Pedagogía*. N.º 148, pp: 29-42.
- MATAMALA, A. (1977): “Influencia de la diferencia de edad interclase en la obtención de calificaciones escolares”. *Infancia y Aprendizaje*. N.º 5, pp: 15-19.
- MATEO ANDRES, J. y RODRIGUEZ ESPINAR, S. (1984): “Precisiones y limitaciones explicativas en los métodos correlacionales. Alternativas metodológicas”. *Revista de Investigación Educativa*. N.º 4. pp: 103-132.
- MIGUEL DIAZ, M. de. (1979): *Diseño de un programa de educación compensatoria en función de los determinantes del rendimiento en el primer ciclo de EGB*. VII Plan de Investigación Educativa. ICE. Universidad de Oviedo.

- MOLINA GARCIA, S. (1984): *Códigos lingüísticos y rendimiento escolar*. Santiago de Compostela: VIII Congreso Nacional de Pedagogía.
- MOLINA GARCIA, S. y GARCIA PASCUAL, E. (1984): *El éxito y el fracaso escolar en la EGB*. Laia. Barcelona.
- MORALES, J. F. (1982): "La investigación sobre el rendimiento académico de la IEA". *Temas de Investigación Educativa*. 2. MEC. Madrid.
- MUÑOZ ARROYO, A. (1977): *Valoración del rendimiento de Centros Docentes de Educación General Básica*. VI Plan de Investigación. ICE. Univers. de Extremadura.
- ORPI GARCIA y PORCAR BECKER, I. (1983): "La inteligencia: un estudio comparativo de las aptitudes cognoscitivas y el rendimiento escolar en adolescentes". *Cuadernos de Psicología*. 1983, II. pp: 15-25.
- PACHECO DEL OLMO, B. y CABALLERO, A. (1977): *Diagnóstico del rendimiento pedagógico en EGB*. Miñón. Valladolid.
- PALOMINO LOPEZ, A. (1979): "La predicción del rendimiento y sus posibilidades orientadoras", en "Dimensiones Psicopedagógicas y predicción del rendimiento escolar". *Revista de Psicología General y Aplicada*. Vol. 34. N.º 160-161. pp: 907-927.
- PEINADO ALTABLE. (1980): *influencia de la relación maestro-alumno en el fracaso escolar, en la Enseñanza General Básica*. IX Plan de Investigación. ICE. Univers. de Valladolid.
- PELECHANO, V. (1977): *Personalidad, inteligencia, motivación y rendimiento académico en BUP*. Informe Final. ICE. Univers, de La Laguna.
- PEREZ GUEDE, A., SANCHEZ TURET, M. y GRAS TORNERO, M. (1979): "El rendimiento escolar del niño hipercinético", en "Dimensiones psicopedagógicas y predicción del rendimiento escolar". *Revista de Psicología General y Aplicada*. Vol. 34. N.º 160-161. pp: 907-927.
- PEREZ SERRANO, G. (1981): *Origen social y rendimiento escolar*. Centro de Investigaciones Sociológicas. Madrid.
- PRIETO ADANEZ, G. y otros. (1982): *La universidad a través de sus alumnos*. Ediciones Universidad de Salamanca. ICE. Salamanca.
- PRIETO, G. y CARRO, J. (1981): "Motivación y éxito académico en la Universidad". *Studia Pedagógica*. N.º 7. pp: 55-61.
- RECARTE, M. A. (1983): "Exito-fracaso escolar al final de la EGB". *Infancia y Aprendizaje*. N.º 23. pp: 23-41.
- REPETTO TALAVERA, E. (1984): *Pluralismo individual y rendimiento académico: un estudio quasi-experimental en 8.º de EGB*. Santiago de Compostela: VIII Congreso Nacional de Pedagogía.
- REPETTO TALAVERA, E. (1984): "Inteligencia, personalidad y rendimiento académico: un análisis de correlación canónica". *Revista Española de Pedagogía*. Año XLII. N.º 166. Octubre-Diciembre. pp: 501-539.

- RIVAS MARTINEZ, F. y ALDOMAR, C. (1975): "Elaboración de un test Hipotético-Inductivo (HI) para la predicción del rendimiento académico en la Universidad Politécnica de Valencia". *Revista de Psicología General y Aplicada*. Vol. 30. N.º 133-134. Marzo-Junio. pp: 409-420.
- RODA FDEZ, R. y MORALES, J. F. (1976): "Rendimiento académico e información de los padres sobre el sistema educativo". *Revista de Psicología general y Aplicada*. Vol. 31. n.º 141-142. pp: 1.048-1.049.
- RODRIGUEZ ESPINAR, S. (1982): *Factores de rendimiento escolar*. Oikos-Tau. Barcelona.
- ROGERO ANAYA, A. "Influjo de las técnicas de estudio en el rendimiento académico". Citado por CARABAÑA, J. en *Temas de Investigación Educativa*. 2. MEC. Madrid.
- ROIG IBAÑEZ, J. (1983): "Más en torno a la variable "agrupamiento" en relación con el rendimiento académico". *Bordón*. Mayo-Junio. N.º 248. pp: 279-313.
- ROSENZWINGE, A. (1979): "Condicionamiento ambiental y rendimiento escolar al entrar en la EGB", en "Dimensiones Psicopedagógicas y predicción del rendimiento escolar". *Revista de Psicología General y Aplicada*. Vol. 34. N.º 160-161. pp: 907-927.
- SAMPER, SAMPER Y SOLER. (1982): *Perspectivas psicológicas y sociológicas sobre el fracaso escolar*. ICE. Univ. Autónoma de Barcelona.
- SEISDEDOS, N. y LOPEZ DE ABERASTURI, N. (1981): "Las evaluaciones y una batería de test (en octavo de EGB)". *Revista de Psicología general y Aplicada*. Vol. 36 (2), pp: 255-262.
- SITE (Servicio de Inspección Técnica de Educación): "Análisis de la magnitud del fracaso escolar considerado globalmente y por niveles", en MOLINA GARCIA, S. y GARCIA PASCUAL, E.: *El éxito y el fracaso escolar en la EGB*. Laia. Barcelona.
- SOLER FIERREZ, E. (1980): "Rendimiento de la enseñanza obligatoria en Andalucía". *Revista Española de Pedagogía*.
- SOPEÑA ALCORLO, A. (1984): "Unos perfiles dominantes de éxito y de fracaso escolar. Orientaciones Educativas". *Educadores*. Vol. XXV. N.º 126. Enero-Febrero. pp: 9-23.
- TEJEDOR TEJEDOR, F. J. (1982): "Análisis de ítems en las pruebas de evaluación del rendimiento". *Revista de Investigación Educativa*. N.º 2. pp: 307.
- TOURON FIGUEROA, J. (1981): "La predicción del rendimiento académico de los alumnos universitarios: algunas implicaciones pedagógicas". *Revista Educação e Seleção*. Julio-Diciembre. Fundación Carlos Chagas. São Paulo. Brasil.
- TOURON FIGUEROA, J. y ARRIETA, F. J. (1983): "La utilización del Programa "PAESSO/10" en el cálculo de ecuaciones de regresión múltiple con el método "WHERRY-DOOLITTLE": su aplicación en la predicción del rendimiento académico". *Revista Española de Pedagogía*. N.º 160. Año XLI, Abril-Junio. pp: 295-310.
- TOURON FIGUEROA, J. (1983): "La selectividad y los factores de rendimiento académico en la universidad". (ICE, Univ. de Navarra). Comunicación presentada al IV Centenario de la Universidad de Zaragoza. Ciclo: "La universidad y panorama del estado actual de la Ciencia". Mayo.

TOURON FIGUEROA, J. (1984): "Predictores académicos del rendimiento del Primer Curso de la Licenciatura de Medicina". Comunicación presentada al VIII Congreso Nacional de la Sociedad Española de Educación Médica. Zaragoza.

VILLAR ANGULO, L. M. (1980): *Modelo inductivo de enseñanza-aprendizaje*. ICE. Univ. de Sevilla.

VILLAR ANGULO, L. M. (1984): "Calidad de enseñanza y supervisión instruccional". *Revista de Investigación Educativa*. (para publicar).

# Ponencias

Revista Investigación Educativa - Vol. 3 - N.º 6 - 1985 (P.127-144)

## Posibilidades de la metodología cualitativa VS. Cuantitativa

por

M.ª Teresa Anguera

Una vieja polémica se ha aviado en los últimos años acerca de las posibilidades de las metodologías genéricamente consideradas como cualitativas y cuantitativas, y no siempre con la necesaria serenidad para favorecer una objetividad que pudiera constituirse en premisa básica. Muy al contrario, los partidarios de unas y otras se ha encastillado en sus posiciones, endureciéndolas, y, como consecuencia, imposibilitando un necesario diálogo entre sí.

Si tratamos de aproximarnos a las bases de la confrontación, se hace necesario deslindar los puntos que podrían considerarse como claves en la configuración de los respectivos paradigmas. El propio uso de este término, versátil y lleno de matizaciones (Kuhn, 1962; Masterman, 1970; Caparrós, 1980) confiere un carácter positivo a la discusión (Patton, 1978; Reichardt & Cook, 1979), que podría tender a plantearse desde uno y otro lado de la controversia, o tratando de su posible convergencia lejana si se arbitran medios que hagan viable la consecución de ciertos objetivos.

Pero, ¿qué objetivos? En este III Seminario de Modelos de Investigación Educativa, al adoptar el *rendimiento académico* como tópico central, deberán aportarse nuevos elementos de discusión que permitan dilucidar en este ámbito y otros relacionados las posibilidades de la metodología cualitativa, considerada tradicionalmente como débil, y que ha polarizado las alabanzas y contestaciones más radicales. Nosotros pretendemos, encuadrados en la perspectiva que ofrece este marco, y desde las bases de la confrontación establecida, concretar la aspiración hacia una perspectiva unificada a lo largo de un proceso que culmine en la viabilidad de un análisis adecuado.

### BASES DE LA CONFRONTACIÓN.

En su acepción más extendida y aceptada, «las metodologías cualitativas se refieren a procedimientos de investigación que dan lugar a datos descriptivos...»

(Bogdan & Taylor, 1975, p. 4.). Esta afirmación, sin embargo, comporta implícitamente un trasfondo que se configuró en la década de los setenta a partir de la cristalización de la perspectiva cualitativa/humanista comenzada en los años sesenta, y que en la actualidad se halla en fase de depuración –no exenta de una sofisticación probablemente exagerada– que permite pensar claramente en su complementariedad con una metodología cuantitativa, a la que incluso puede llegar a superar en algunos casos en grado de formalización (Haberman, 1978, 1979).

Si atendemos a sus raíces lejanas, Dilthey y Rikert plantearon la polémica en el siglo XIX al recabar la originalidad metodológica/sustantiva de las ciencias humanas frente a las ciencias naturales, y que dependía de la naturaleza de su objeto de estudio. La Escuela de Chicago, dominante en el período de 1915-1935, logró que, a través de Znaniecki, se impusiera con fuerza la perspectiva humanista/cualitativa: Comprensión frente a explicación, y ciencias idiográficas frente a las nomotéticas, constituyan los dos polos o ejes básicos, a los cuales pronto se añadiría la idea de calidad/cantidad, con su referencia inmediata a la medición y uso o no de métodos estadístico-matemáticos, y con lo que, a su vez, ello pudiera comportar respecto a datos «duros» y «blandos» (Alvira, 1983). En la época de la Escuela de Chicago, en donde se utilizaban los documentos personales (cartas o epístolas), entrevistas y biografías, como forma de recoger y acumular datos –basta recordar la famosa obra *El campesino polaco*, de Znaniecki–, éstos eran evidentemente «subjetivos» en tanto en cuanto expresaban el punto de vista del autor, y en ellos, en palabras de Madge (1963, p. 74), «... el material se utilizó para el análisis cualitativo y no cuantitativo. Se sigue inevitablemente del método como se recogieron las cartas que no hay base para la representatividad, ni análisis estadístico, ni verificación de hipótesis, y con ello se fraguaba en la década de los treinta la idea de que los datos y análisis cualitativos cumplen un importante papel en las fases preparatorias de las investigaciones, así como en los estudios exploratorios, sin que en ningún caso fueran utilizables en fases confirmatorias. Esta idea se radicaliza en la siguiente década, en que se afirmaría tajantemente la imposibilidad de llegar a la verificación rigurosa sólo mediante datos o análisis cualitativos.

Según Alvira (1983), el cambio radical que empezó a plantearse en los años sesenta tiene su origen en la crisis del justificacionismo y en la tajante separación entre contexto de descubrimiento y de justificación en la filosofía y metodología de las ciencias, por una parte, y en el avance tecnológico en la medición, tratamiento y análisis matemático de los datos, por otra. En el primer caso (Knorr-Cetina, 1981) se cuestiona la posibilidad de verificación de teorías, por lo que la construcción de éstas de forma adecuada pasó a ocupar un punto crucial; y en el segundo, es evidente que lo cualitativo se puede transformar en cuantitativo, por lo que dejan de tener sentido las barreras tradicionales y se puede tender a la complementariedad, aunque con matizaciones.

En la actualidad, el número de dimensiones/ejes en que se fundamenta la pugna metodológica ha desbordado las previsiones iniciales, extendiéndose a as-

pectos muy numerosos y variados que afectan a la epistemología, la metodología y las técnicas de investigación. De ahí la denominación de paradigma, que, a pesar de su indefinición (Masterman, 1970), aquí nos lleva a saber la existencia de una visión del mundo compartida por un grupo de científicos que implica explícitamente una metodología específica, sea cualitativa o cuantitativa, caracterizada por problemas, procedimientos, técnicas, vocabulario y "tendencia interpretativa". Como producto, pues, de la actividad y producciones de comunidades de científicos, surgieron verdaderos "manifestos" con direcciones opuestas.

Por una parte, Campbell & Stanley (1966) y Riecken & Boruch (1974), con diferencias no relevantes, abogan a favor de los métodos cuantitativos en tanto en cuanto consideran el experimento como el método que, por excelencia, permite obtener información fidedigna y válida, que, por supuesto, va a someterse a los mandatos propios del diseño que corresponda en cada caso. Desde la posición opuesta, Weiss & Rein (1972), Parlett & Hamilton (1976) y Guba (1978) denuncian el carácter restrictivo propio de las tradiciones experimentales y psicométricas dominantes en la investigación educativa. Entre ambas posturas se sitúa el posterior trabajo de Cook, que, con Campbell (Cook & Campbell, 1976), realiza una defensa de los diseños cuasi-experimentales, a la vez que revaloriza la perspectiva cualitativa como necesario e insustituible complemento de la cuantitativa.

Sería poco fructífero, sin embargo, centrar la discusión en el comentario de adjetivos o listados de dimensiones que han sistematizado diversos autores, como Filstead (1970), Bruyn (1972), Halfpenny (1979), Reichardt & Cook (1980), y que prácticamente convergen en caracterizar lo cualitativo por ser flexible, exploratorio, inductivo, holístico, interpretativo, fenomenológico, relativista, idiográfico y basado en la realidad, mientras que a lo cuantitativo se le asignarían los adjetivos opuestos. No vemos tampoco a abogar por los métodos cualitativos al considerarlos "mejores" que los cuantitativos, sino a deslindar diferentes niveles de discusión que en la práctica han estado sujetos a colisiones continuadas y han introducido con ello un considerable grado de confusión. Y estos posibles niveles son el epistemológico y el metodológico/técnico.

Algunos autores (Alvira, 1982), que asimilan el nivel epistemológico a la consideración paradigmática, están de acuerdo en que los evaluadores no se sienten libres para cambiar su postura paradigmática cuando sea necesario, lo cual actúa de pesado lastre que impide la adopción de un punto de vista flexible y adaptativo. De aquí deriva la apremiante necesidad de delimitar los puntos de diferenciación a nivel epistemológico o paradigmático si después pretendemos analizarlos metodológicamente:

A) En la actualidad, la opción deliberada por el descubrimiento de la teoría a partir de los datos o "teoría fundamentada" (Glaser & Strauss, 1968; Turner, 1981) se apoya en una revisión de la inducción analítica. El objetivo es la comprensión precisa del fenómeno a estudiar, y desde este punto de vista, la inducción analítica se trataría de un tipo de análisis causal, ya que suministraría las necesarias, pero no suficientes, condiciones para la presencia de tal fenóme-

no; como indicaba Robinson (1951, p. 814), “el método permite estudiar sólo aquellos casos en los que el fenómeno ocurre, y no aquellos en los que no ocurre. Estudiar casos en los que el fenómeno no ocurre nos involucraría en una inducción enumerativa”. Ahora bien, el propio Robinson muestra como “en la práctica” la inducción analítica es similar a la enumerativa, puesto que, incluso en el caso que se utilizaran análisis sofisticados (lo cual, por otra parte, aquí no es previsible), las variables no cumplirían ni las condiciones necesarias ni suficientes en el estricto sentido lógico de estos términos (Miller, 1982).

Cada fenómeno presenta un número muy elevado de características, y de ahí la necesidad de llegar al desarrollo de la “teoría fundamentada” a partir de una serie de estadios (Turner, 1981) que se inician con la formación de categorías, y siguen con su saturación y definición depuradas, para poder detectar y considerar nexos entre ellas, los cuales mostrarán las conexiones existentes que se sitúan en la base de la teoría.

B) Se ha contrapuesto vivamente la prioridad en asegurar que un estudio sea interna o externamente válido. Precisamente uno de los aspectos más importantes que tiene planteada la investigación “experimental” consiste en la obtención de inferencias válidas sobre la relación de variables, basadas en cuidadosas observaciones en la recogida de datos, y de ahí la necesidad del control adecuado; pero además, una vez el estudio está perfectamente diseñado, debe poder ser generalizado el resultado obtenido. Éste es el difícil equilibrio entre la exigencia del control y la necesidad de descubrir relaciones constantes y de gran alcance entre las variables, y en donde se halla la clave sobre la validez. Ante la pregunta sobre cuál debe prevalecer sobre la otra, la respuesta –una vez más– está en función de la posición de partida: Para la metodología cuantitativa interesa conseguir un máximo grado de validez interna, acompañada del razonamiento de que si falta un control adecuado en la relación causal simple o compleja entre variables, mal podríamos generalizar estos resultados incorrectos, producidos por variables extrañas contaminadoras; en este sentido, Campbell y Stanley (1973, p. 16), afirman: “Mientras que los problemas de la validez interna pueden ser solucionados, dentro de los límites de la lógica de la estadística probabilística, los problemas de la validez externa no son solucionables de un modo lógico, de ninguna manera tajante y clara. La generalización implica siempre la extrapolación a una realidad no presente en la muestra estudiada. Esta extrapolación se realiza bajo el supuesto de que se conocen las leyes adecuadas... lógicamente no podemos generalizar en absoluto”.

Una manera de resolver este problema ha sido apelar a la repetición de las investigaciones, de modo que la prueba de su validez externa (Birnbrauer, 1981) venga dada por la replicación de los hallazgos por diferentes investigadores, en diferentes situaciones y con diferentes poblaciones. Y en cualquier caso entendemos que no tendría sentido referirnos a esta distinción si sólo se tratara de estudios *descriptivos*, si no que es en los *explicativos* donde se plantean los problemas de la validez interna y externa en toda su extensión, lo cual es lógico si pensamos que constituyen el único marco para el establecimiento de las relaciones de

causalidad, y partiendo de la base de que el investigador elige las variables explicativas de acuerdo con las teorías científicas sustantivas, es decir, no es el método cuantitativo por si mismo, sino el sistema teórico previo, el que le llevará a decir cuáles son variables y cómo deberán ser tenidas en cuenta.

En la vertiente opuesta a la de Campbell & Stanley (1966), Brunswik (1952, 1956) piensa que las condiciones experimentales deberían representar una muestra aleatoria del número de condiciones al que posteriormente van a extenderse los resultados. De esta forma, mediante una adecuada selección de las variables, se evita la artificialidad de la experimentación. La utilización, por parte de Brunswik, del término "diseño representativo", como opuesto al de "diseño sistemático", no significa otra cosa que la transposición del principio estadístico del muestreo de los sujetos al muestreo de las situaciones y condiciones experimentales (Hursch, Hammond & Hursch, 1964; Snow, 1968). Lo que en definitiva propugna Brunswik es que se aplique el muestreo a las diferentes situaciones estimulativas (en lugar de ser poblacional), es decir, en relación con su generalidad ecológica, discutida por Bracht & Glass (1968).

En este sentido, la investigación educativa en general, y en el ámbito del rendimiento en particular, según Snow (1979), no parece haberse llegado a una teoría consolidada ni a generalizaciones lo suficientemente fuertes para superar variaciones aparentemente pequeñas en alumnos, profesores, situaciones de enseñanza, material a aprender, postests usados como criterio, etc. Sin embargo, y por esta misma razón, es conveniente preguntarse si se da la compatibilidad entre el objetivo y la metodología utilizadas (Kruglanski & Kroy, 1976). Brunswik diría que deberíamos seguir usando el diseño sistemático para aquello para lo que es apropiado, pero hasta el momento apenas se ha logrado ajustar esa metodología al funcionamiento adaptativo o probabilista de la conducta humana, ni derivar descripciones molares de esa conducta en situaciones naturales multidimensionales, como las de valoración del rendimiento, y máxime teniendo en cuenta que la mayoría de los experimentos educativos no son externamente válidos, es decir, no son representativos.

C) La discusión entre lo cualitativo y lo cuantitativo a nivel epistemológico pasa también por la consideración, ya apuntada, de naturalidad frente a artificialidad de la situación, con lo cual nos acercamos al segundo nivel de discusión. Es evidente la existencia de diferentes procedimientos que se sitúan en todo el arco de posiciones existente entre ambos extremos, y por su carácter de bisagra y lo que ello comporte de clarificador, es interesante analizar y tener en cuenta el papel jugado por la quasi-experimentación. Al hallarse ésta diferenciada de los experimentos y de las situaciones naturales, aunque más escorada a los primeros, no es extraño que en recientes discusiones se hubieran considerado como puntos de apoyo o pilares básicos la objetividad y la posibilidad de verificación y falsación.

Por lo que se refiere a la objetividad, existe un punto de inflexión, dado que en la investigación cuantitativa se supone que es máxima, y mínima en la cualitativa. Morgan (1983) afirma que los quasi-experimentalistas creen que sus ob-

servaciones están impregnadas no de una única teoría, sino de muchas, y tienen fe en datos que han sido obtenidos con fiabilidad suficiente después de repetir la medida; esta concepción iguala la objetividad con la fiabilidad intersubjetiva más que con la medida infalible de lo externo. Esta suposición de que la objetividad es sólo factible como verificabilidad crítica intersubjetiva sugiere por qué el diseño cuasi-experimental se ha asociado con el operacionalismo múltiple (Cook & Campbell, 1979). Los cuasi-experimentalistas son isomórficos con las medidas, y ven a éstas como conteniendo una única variancia que no forma parte del constructo que se examina. De aquí surgirán las propuestas de triangulación, para utilizar medidas que reparten la variancia supuesta al representar el constructo de interés, pero que tienen fuentes únicas de variancia teóricamente irrelevante. La misma lógica podría considerarse referida a la replicación.

Y en segundo lugar, los cuasi-experimentos permiten inferir, y la inferencia depende lógicamente de una información imperfecta. Los cuasi-experimentos sirven para "probar" relaciones causales, aunque nunca éstas estén "probadas" mediante cuasi-experimentos; en efecto, la covariación de dos medidas no comprueba la existencia de una conexión causal debido a la fiabilidad lógica de existencia de otras hipótesis a las que también pueden ajustarse los datos. De aquí la importancia que tiene el que las teorías permitan predicciones numéricamente precisas o bien el que se obtengan resultados únicos. Por tanto, la perspectiva cuantitativista se presenta como justificaciónista dentro del contexto de la verificación, mientras que el contexto del descubrimiento quedaría bien para la suerte o intuición, o bien para la perspectiva cualitativa (tal como se ha entendido tradicionalmente). Desde la misma lógica, la metodología cualitativa se ha presentado sucesivamente como: a) Un esfuerzo preliminar exploratorio que antecede a la investigación cuantitativa proporcionándole ideas, hipótesis, conceptos; b) una forma determinada de descubrir, generar y construir teoría (conceptos, clasificaciones, hipótesis) enraizada e imbricada en la realidad (teoría fundamentada o "grounded theory"); c) una forma más de verificar, contrastar o justificar teorías e hipótesis. Autores de prestigio reconocido, como Lazarsfeld (1955), afirman que los datos cualitativos, por su riqueza descriptiva, juegan un papel capital en las fases exploratorias de la investigación, no pudiéndose pretender una verificación rigurosa de tipo estadístico o experimental, sino solamente lo que se podría llamar "confirmación" o "corroboration". Sin embargo, Lazarsfeld, al no adoptar una postura metodológica/normativista, es decir, al describir y analizar lo que se hace, se ve obligado a reconocer (Alvira, 1983) que el análisis cualitativo se utiliza para la verificación o prueba, aunque no se atreve a igualarla a la prueba cuantitativa o estadística.

## ASPIRACIÓN HACIA UNA PERSPECTIVA UNIFICADA

Desde la primera consideración epistemológica, serían muchas más las ventajas que los aspectos negativos si se tendiera a una perspectiva unificada del co-

nocimiento cualitativo y cuantitativo, consistente con el estado actual en que se halla la filosofía de la ciencia.

En principio, debería intentarse una redefinición del debate, eliminando la antes necesaria elección de paradigmas, dado que éstos se han configurado por una serie de elementos concatenados por tradición, por lo que histórica, pero no lógicamente, quedaron establecidos. Un investigador no tiene por qué adherirse ciegamente a uno de ambos paradigmas, sino que puede elegir libremente una relación de atributos que indistintamente provengan de uno u otro si así se logra una adaptación flexible a su problemática. Y aquí queremos dejar constancia de forma clara y definida de que la necesaria separación entre el nivel epistemológico y el metodológico/técnico nos lleva a la afirmación de que los métodos –entendidos por supuesto desde la aceptación de metodologías específicas– pueden utilizarse conjuntamente según las exigencias de la situación investigadora (Reichardt & Cook, 1979), no confundiéndose con los razonamientos alternativos que dieron lugar a ambos paradigmas.

Para salir del bloqueo que puede provocar esta situación, es necesario caracterizar la forma como diversas técnicas de recogida de datos que han sido utilizadas, con el fin de ensanchar sus posibilidades y contemplar entrecruzamientos enriquecedores que podrían contribuir lentamente a la perspectiva unificada a la que aspiramos.

En este sentido, Erickson (1977) afirma que lo que hace mejor y más esencialmente la metodología cualitativa es describir incidentes clave en términos descriptivos y de forma funcionalmente relevante, contextualizándolos en función de un ámbito más amplio. Si lo analizamos friamente, la idea básica que guía la investigación cualitativa se puede reducir a la utilización exclusiva de escalas nominales, lo cual equivale a someter la situación a un necesario proceso de deslinde de unidades, elaboración de un sistema de categorías coherente, y codificarlas si es necesario; para llevar a cabo esta serie de operaciones no será necesario más que una progresiva reducción de datos –que no de información–, y los puntos de partida pueden ser diversos (observación directa, observación indirecta, estudio de casos, entrevista,...), tendiendo a comportar una no intervención o incidencia del investigador en la situación, que en todo caso puede desempeñar a lo sumo el papel de entrevistador u observador participante. Queda claro, sin embargo, que el investigador que utiliza métodos cualitativos no es necesariamente un fenomenólogo, ni tiene porque ser subjetivo el registro obtenido, ni ser exploratorio, ni proceder por vía inductiva, ni llegar a falta de acuerdo entre diferentes observadores, ni ser forzosamente idiográfico, ni siempre holístico, ni hallarse reducido a algo puntual, ...

Por otra parte, los investigadores cuantitativos tienden a traducir sus observaciones en números, y estos valores numéricos se asignan por conteo y medida (Argyris, 1980), lo cual permitirá descubrir, verificar o identificar relaciones causales entre conceptos que derivan de un esquema teórico elaborado “*a priori*”. Habitualmente, pues, se sigue una vía deductiva, y para llevar a cabo el contraste de la hipótesis será preciso cumplir el requisito de representatividad y

aleatorización, lo cual comportará unas adecuadas técnicas de muestreo y diversos procedimientos que garanticen el necesario control, así como definiciones operacionales que harán factibles sofisticadas técnicas de recogida de datos y su posterior análisis.

Los dos tipos de conocimiento pueden beneficiarse mutuamente entre sí, y la ciencia normal los utiliza de manera conjunta en muchas ocasiones, dando garantías de su posible complementariedad, que en ocasiones presenta ciertamente graves problemas por su costo en tiempo y dinero, o bien por falta de personal preparado al efecto, pero que en cualquier caso se trata de superar ambas tradiciones (Von Wright, 1979).

## CONFIGURACIÓN DE UN «DISEÑO CUALITATIVO»

La investigación cualitativa en el ámbito educativo, y, por tanto, en el del rendimiento académico, se halla en una etapa de gran vitalidad, habiéndose ya incorporado una cantidad considerable de términos que forman parte de un nuevo léxico que está siendo bien aceptado en general por quienes han estado utilizando únicamente métodos cuantitativos. Además, los investigadores cualitativos han empezado a salir de su letargo y abandonar su antigua postura defensiva, que en ningún momento tenía que venir marcada por la naturaleza de los datos, dado que el registro puede ser tanto o más riguroso que en estudios correlacionales o experimentales. Actualmente, el principal interés se ha desplazado a una mayor responsabilidad en su utilización y un estudio deliberado de las particularidades que deben cumplirse en la correspondiente investigación cualitativa: cómo se lleva a cabo la recogida, cómo se almacenan, particionan y analizan los datos, y cómo podrán integrarse los resultados obtenidos, con lo que dejará de experimentarse una cierta sensación de vacío (Huberman & Miles, 1983; Miles & Huberman, 1984).

De cara a la configuración de un “diseño cualitativo” ejerce un peso indudable la falta de cánones de procedimiento o reglas específicas de decisión para el análisis de los datos (Huberman, 1981), así como una cierta imprecisión en la medida, debilidad en la generalización, cierta vulnerabilidad, y dudas acerca de la forma cómo debe llevarse a cabo la consideración del contexto (Miles, 1979). En el ámbito del rendimiento académico, así como en cualquier otro, se impone una primera tarea de reducción de datos, dado que éstos no se analizan inmediatamente después del registro por lo general, sino que se almacenan progresivamente a medida que avanza el proceso de recogida de datos, y precisamente la elección de la estrategia de reducción estará muy relacionada o incluso determinará el tipo de análisis posterior.

La reducción de datos se refiere al proceso de seleccionar, focalizar, simplificar, extractar y transformar los datos brutos que aparecen en las notas tomadas en el registro. Por lo tanto, no es algo separado del análisis, sino que es parte de él en la medida en que los “clusters” y particiones deberán obedecer a criterios

fijados sobre qué informaciones conviene tener en cuenta simultáneamente y qué interrogantes hay que responder. Por reducción de datos no necesariamente entendemos cuantificación, ya que puede llevarse a cabo mediante un organigrama, u ordenación, etc., con lo que el texto narrativo se transformaría convenientemente, facilitando su evaluación (Wachtman, 1982) al estructurarse mediante análisis de contenido (Woodrum, 1984), que, a su vez, hará posible la conexión entre el sustrato teórico y el material empírico.

Debido precisamente a que la reducción de datos forma parte del análisis, es importante recordar el concepto de "triangulación" (Denzin, 1970; Webb, 1978), dado que es frecuente la obtención de medidas o indicadores simultáneos o múltiples acerca del mismo atributo o constructo –perfectamente factible en el caso del rendimiento-. Con sólo un método es imposible separar el sesgo del método de la cantidad o calidad subyacentes que se quiera medir. Pero se pueden utilizar varios métodos conjuntamente para "triangular" la verdad subyacente. Como indican Reichardt & Cook (1979), aunque dos o más métodos cualesquiera pueden utilizarse para este propósito, es mejor que sean dispares, ya que sus sesgos serán heterogéneos, y de aquí la conveniencia de que converjan los métodos cualitativos y cuantitativos, que, además, pueden sugerir nuevas vías de actuación.

Si desde la perspectiva cuantitativa el investigador se separa de la "sustancia" que configura la materia de estudio con el fin de descubrir las regularidades que posibilitan la predicción, y desde una perspectiva cualitativa se enfatiza la interpretación, en contacto con la propia realidad (Smith, 1983), será útil, efectivamente, la complementación de ambas. Los datos recogidos mediante la metodología cualitativa permitirán al evaluador que las situaciones seleccionadas no estén restringidas por categorías de análisis predeterminadas, pudiendo así proporcionar mayor profundización y detalle (Patton, 1984); y los métodos cuantitativos, por el contrario, requieren la utilización de estímulos estandarizados, de forma que todos los datos obedezcan a ciertas categorías de respuesta, pero con la ventaja de facilitar la comparación y agrupación de los datos mediante técnicas estadísticas.

En consecuencia, los datos cualitativos consisten en descripciones detalladas de situaciones, eventos, sujetos, interacciones y conductas observadas; y, si la observación es indirecta, se puede también obtener datos de documentos, correspondencia, historias de casos, etc., así como material procedente de experiencias, actitudes, creencias y pensamientos, en el caso de la auto-observación. Es innegable que pueden proporcionar profundidad y detalle, como se ha comentado anteriormente, ya que emergen los datos de una descripción cuidadosa, y en cualquier caso dependerán del objetivo de cada estudio particular.

Lofland (1971) sugiere la existencia de cuatro elementos en la recogida de datos cualitativos. El evaluador debe sentirse cercano a la situación para comprenderla con exactitud, debe dejar constancia de todos los hechos percibidos (es decir, conductas manifiestas), llevar a cabo una descripción pura (no evaluativa), y registrar convenientemente. Se trata de datos que se han considerado tradicio-

nalmente como “blandos” (Weimer, 1979; Guba & Lincoln, 1982), pero ello no significa que se sigan evaluando programas educativos a espaldas de una metodología científica (Campbell, 1979), sino que, por el contrario, en ambas metodologías (Ianni & Orr, 1979) deberá procederse a la delimitación de unidades.

En la evaluación del rendimiento, ¿puede verse la situación como un “continuum” o como una secuencia de unidades discretas? La búsqueda de unidades, su organización y su validación empírica constituye un importante problema a resolver, y en donde la dificultad metodológica se halla en la decisión del criterio (Condon & Ogston, 1967), que tendrá también una evidente repercusión en los parámetros obtenidos; además debe tenerse en cuenta la posibilidad de diversos niveles de descripción (Coll, 1981), que si bien resultaría poco útil si es ética por la profundidad que ello reportaría, en cambio, como descripción émica se logra que se identifiquen y precisen las categorías significantes, para lo cual existe un evidente apoyo en los indicadores empíricos de éstas y de su articulación, tal como aparecen precisadas en los modelos explicativos. La construcción de los sistemas de descriptores se puede considerar unidireccional: comportamiento del sujeto en situación de evaluación del rendimiento → corpus de datos → plantilla de análisis del observador → descripción ética o de primer nivel (registro y codificación de los datos) → análisis de las relaciones internas de los datos tal como han sido codificados → descripción émica o de segundo nivel (categorías significantes y funcionales) (Coll, 1981). Puede darse esta posibilidad correspondiente a una vía inductiva, o bien cuando el investigador parte de un determinado sistema, y asimismo, o bien se acepta que los criterios que guían al investigador son impuestos por los propios datos, o bien aceptamos que vengan dados por sus categorías significantes.

En definitiva, el desarrollo del proceso de categorización permite determinar el entorno en la forma que es relevante a una especie, sea cual sea su nivel de complejidad (Sáinz, 1984), y en el ámbito del rendimiento académico, en la índole misma de las categorías o de los conceptos habituales deben estar dadas tanto las condiciones de su mútua articulación como aquéllas en que pueden adaptarse a ciertas situaciones específicas. Existe una distinción meramente funcional entre categorización y conceptualización, y se justifica por un proceso común subyacente que se cualifica en función de la relevancia del proceso que se juzga y según cuáles sean los sistemas que lo conforman (Mayor y Sáinz, 1984), a la vez que hace posible que el hombre se adapte a un número indefinido de situaciones reordenando las descripciones del entorno. La conceptualización, que presenta un carácter especializado, deberá determinar los aspectos relevantes a una secuencia y la organización óptima que, según el estado del sistema, es pertinente a la previsión de los acontecimientos, por lo que, aunque se puede derivar cómo opera el proceso de categorización a partir de la conceptualización, hay que tener presente que todo modelo de categorización lo es primero de conceptualización. Son diversos y múltiples los diferentes modelos de categorización, y en cualquier caso se precisa una segunda distinción en el análisis de la estructura de la categoría, en donde los conceptos de núcleo conceptual y proce-

dimiento de interpretación expresan, respectivamente, la descripción unitaria que permite reconocer los miembros de una categoría, y el procedimiento por el que una categoría específica se asigna a una instancia particular. El núcleo incluye los atributos abstractos primariamente responsables de ciertas relaciones entre conceptos, en tanto que los procedimientos de identificación determinan los rasgos que son relevantes en la categorización de los acontecimientos del mundo real susceptibles de ser tratados como miembros del concepto (Smith & Medin, 1981). En cada caso diferirá la relativa complejidad en función de la situación estudiada, que influirá en el carácter molar, molecular, o mixto del sistema de categorías, tanto si éstas proceden de observación directa, como de análisis de contenido, como de una entrevista, etc., posibilitando imponer escalas y dimensiones en datos cualitativos (Patton, 1980) que, en cualquier caso, proceden de situaciones espontáneas o planificadas en clase, en donde puede tomarse el rendimiento académico de un alumno en determinada habilidad como base de evaluación (Adams, 1983).

Es frecuente que la recogida de datos cualitativos implique la utilización de categorías multidimensionales, de forma similar a cómo un evento cualquiera tuviera que ser incorporado a un registro descriptivo sistemático en donde tuviera que conceptualizarse como proceso –si se consideraran sus antecedentes, duraciones, contexto, consecuencias–, o bien como marcador o producto si señala una transición o cambio ocurrido. Desde un punto de vista mecanicista, el tiempo se consideraría como atomístico, formado por elementos discretos, mientras que el proceso se constituiría por una sucesión de actividades en un contexto que incluyen el paso del tiempo. Esta confusión proceso-producto (Reese & Smyer, 1983) deberá resolverse mediante una adecuada dimensionalización de tales eventos, a lo cual coadyuva la frecuente existencia de una dimensión “saliente” o notoria que permitiría ser tratado funcionalmente como unidimensional, si bien aquí la discusión se encaminaría a considerar o rechazar su posible carácter normativo, lo cual tendrá importantes implicaciones (Datan, 1983).

## VIABILIDAD DE ANÁLISIS DESDE LA INVESTIGACIÓN CUALITATIVA

Por todo lo comentado anteriormente, es evidente que existe un posible paralelismo entre métodos cualitativos y cuantitativos, debiendo refutarse la oscilación pendular de uno a otro extremo detectada hasta el momento presente. Y si es cierto que el excesivo énfasis que se observa actualmente sobre los métodos cuantitativos sólo podrá corregirse por otro de igual intensidad pero opuesto que se ha dado a los métodos cualitativos, entendemos –salvando el radicalismo que podría atisbarse en esta afirmación– que es imprescindible realizar urgentemente una labor de explotación y previa exploración de posibilidades de análisis de datos que son cualitativos.

Recordando que “los datos cualitativos consisten en *descripciones detalladas*”

das de situaciones, eventos, sujetos, interacciones y conductas observadas; *citas directas* de sujetos acerca de sus experiencias, actitudes, creencias y pensamientos; y fragmentos o pasajes enteros de documentos, correspondencia, registros e historias de casos" (Patton, 1980, p. 22), es evidente que darán lugar a registros narrativos sin una predeterminación referente a categorías estandarizadas como las respuestas múltiples preparadas con relación a cuestionarios o tests. El evaluador que sigue métodos cualitativos pretende captar con el máximo rigor lo que los sujetos hacen o dicen (Anguera, en prensa); es decir, trata de describir la realidad, que se materializa en un continuado devenir o flujo de conducta. En consecuencia, la tarea para el metodólogo cualitativo es la de lograr un marco en cuyo seno los sujetos puedan responder de forma que sean fieles con precisión a su experiencia, y, sin caer en un reduccionismo absurdo, proceder a una sistematización mediante la categorización y dimensionalización indicadas, lo cual posibilitaría una cierta cuantificación del conocimiento (Kvale, 1983), que no equivale a afirmar que deberá llevarse a cabo una cuantificación de tal conocimiento, y, en este sentido, manifestamos nuestras reservas a las palabras de Mussen, Conger & Kagan (1979, p. 13), según las cuales "la medida en que las observaciones pueden ser cuantificadas (trasladadas a números) es frecuentemente un buen índice de la madurez de la ciencia". Por supuesto, la cuantificación no es criterio absoluto de ciencia, y, por el contrario, defendemos que, desde la investigación cualitativa, se produzca una integración de aproximaciones cualitativas y cuantitativas respecto al análisis de datos.

De acuerdo con Miles & Huberman (1984), en la investigación cualitativa los números tienden a ignorarse, pero, sin embargo, no podemos renunciar al conteo de las frecuencias de acciones, eventos, o categorías por nosotros establecidas, y, además, deberemos en muchos casos constatar su secuenciación. Ello nos permite establecer dos ejes básicos (Anguera, 1985 a, 1985 b), que sustentarían un sistema de coordenadas útil para una sistematización de análisis viables: Carácter idiomático-nomotético del estudio, y puntualidad-temporalidad en la situación. Actualmente se están llevando a cabo estudios parciales dirigidos a la extrapolación de diversos planteamientos específicos, interesando resaltar especialmente dos de los cuadrantes resultantes del cruzamiento de tales ejes:

a) El seguimiento en un estudio idiomático constituye, por esencia, una situación óptima para identificar relaciones contingentes entre datos a través de diferentes posibilidades de concurrencia y secuencialidad. El proceso de recogida de datos se orienta hacia una búsqueda que en la mayor parte de los casos resulta sumamente interesante sobre la existencia de patrones de conducta, los cuales se refieren a la aparición regular de acciones, eventos o comportamientos que se manifiestan ordenados de forma similar, y relacionados con uno de ellos considerado como preeminente por tener un carácter de desencadenante o inicializador de una serie. Estos patrones nos muestran, después de un potente proceso de reducción de datos, la forma de obtener un extracto o esquema de la información registrada y almacenada a lo largo de diferentes sesiones que, con la simplificación de lo accesorio a que se somete, es capaz de mostrar el prototipo

de ordenación secuencial de las conductas básicas que conforman el sistema de categorías utilizado en el registro. Las posibilidades investigadoras que así se abren sobre el rendimiento académico son muy amplias, facilitando un estudio de sujetos individuales con una ausencia o presencia de problemática, y que se hallan inmersos en un determinado contexto.

b) Es necesario llevar a cabo igualmente una depurada adecuación en los casos en que se plantea no un seguimiento secuencial en la generación de conductas, sino un cruzamiento transversal –y por tanto funcional– de datos correspondientes a uno o varios sujetos. El problema que existía tradicionalmente radicaba en la falta de adecuación entre la situación real y las posibilidades de análisis existentes, ya que los procedimientos conocidos como cuantitativos –que han sido los habitualmente utilizados– tienen una zona de cobertura que no incluye gran parte de los tipos de datos procedentes de registros que han dado lugar a variables categóricas como resultantes de la dimensionalización y categorización, especialmente en situaciones complejas. Cubriendo este espacio, en la última década se han desarrollado y expandido con fuerza procedimientos de análisis basados en modelos “log-linear”, que brindan la solución al superar las viejas formas utilizadas en la búsqueda de relaciones existentes entre variables medidas a niveles discretos. Ante la configuración de datos obtenidos cuando bajo diversos criterios se llevan a cabo las correspondientes categorizaciones, el análisis a que se someten se encamina a la comprobación de existencia de un grado de ajuste óptimo entre las frecuencias empíricas observadas (y constatadas en la tabla inicial de clasificaciones cruzadas) y las teóricas obtenidas a partir de los supuestos de diversos modelos teóricos –ordenados jerárquicamente– que se estructuran de forma similar a los del AVAR, aunque a partir de una relación multiplicativa transformada en aditiva mediante la adopción de logaritmos neperianos.

En el estudio del rendimiento académico entendemos que reportan gran utilidad, permitiendo un importante afinamiento a partir de pautas de categorización que presentan gran complejidad (Anguera y Blanco, 1984) al responder a estructuras, procesos o estilos cognitivos del sujeto, que en muchos casos habrá que inferir a partir de la observación sistemática en la ejecución de diferentes estrategias o en las distintas etapas de una toma de decisiones; además, presenta una versatilidad suficiente para el estudio de tablas de contingencia de alta dimensionalización, por lo que actualmente ya no es un problema insoluble el análisis serio y tratamiento riguroso de situaciones caracterizadas por el carácter cualitativo de sus datos.

Por supuesto, aquí no se agotan en absoluto las posibilidades de análisis de datos cualitativos, sino que hemos tan solo tipificado dos de las situaciones que presentan un mayor interés. Muchas más serían extrapolables desde planteamientos cuantitativos; éstas son, por esencia, propias de investigaciones cualitativas.

## EPÍLOGO

Es mucho lo que podría discutirse acerca de las ventajas de ambas metodologías, tratando cada una de ellas de ganar un espacio milimétrico a su alternativa, pero la dicotomización es mala por naturaleza en Ciencias Humanas, y el reto real es adaptar sin perjuicios los métodos de investigación al problema del rendimiento académico. De aquí la búsqueda de su complementariedad, tarea por otra parte árdua si partimos de la escisión que siempre ha existido, pero posible de inmediato a nivel de método lógico, y esperamos que en un futuro inmediato a nivel epistemológico/paradigmático. En el debate entre las metodologías cualitativas y cuantitativas no hay, pues, ningún perdedor, y sí un solo ganador: la voluntad de acercamiento, que ya ha empezado a producir algunos frutos.

## REFERENCIAS BIBLIOGRÁFICAS

- ADAMS, G.S., (1983) *Medición y evaluación en Educación, Psicología y «Guidance»*. Barcelona: Herder.
- ALVIRA, F., (1982) La perspectiva cualitativa y cuantitativa en las investigaciones sociales. *Estudios de Psicología*, 11, 34-39.
- ALVIRA, F., (1983) Perspectiva cualitativa-perspectiva cuantitativa en la metodología sociológica. *Revista Española de Investigaciones Sociológicas*, 22, 53-75.
- ANGUERA, M.T., (1985 a) Directrices básicas en el análisis de datos observacionales. En M.T., Anguera. *Metodología de la observación en las Ciencias Humanas* (3.<sup>a</sup> edic. ampliada). Madrid: Cátedra.
- ANGUERA, M.T., (1985 b, Junio) *Investigación naturalista: Planteamiento a partir de la metodología observacional*. Trabajo presentado a la 1.<sup>a</sup> Reunión Anual de la «Societat Catalana de Recerca i Teràpia del Comportament», Mesa Redonda de Investigación Básica. Sitges (Barcelona).
- ANGUERA, M.T., (en prensa) Observación. En S., Molina (Dir.) *Diccionario temático de Educación Especial*. Madrid: C.E.P.E.
- ANGUERA, M.T., y BLANCO, A., (1984) *Análisis de datos cualitativos en Psicología Cognitiva: Sistemas alternativos de cuantificación y evaluación en el marco de una estricta metodología científica*. Proyecto subvencionado por la Comisión Asesora de Investigación Científica y Técnica (CAICYT). Barcelona: Departamento de Psicología Experimental.
- ARGYRIS, C., (1980) *Inner contradictions of rigorous research*. New York: Academic Press.
- BIRNBRAUER, J.S., (1981) External validity and experimental investigation of individual behaviour. *Analysis and Intervention in Developmental Disabilities*, 1, 117-132.

- BOGDAN, R., & TAYLOR, S.J., (1975) *Introduction to qualitative research methods*. New York: Wiley & Sons.
- BRACHT, G.H., & GLASS, G.V., (1968) The external validity of experiments. *American Educational Research Journal*, 5, (4), 437-474.
- BRUNSWIK, E., (1952) The conceptual framework of Psychology. *International Encyclopedia of Unified Science*, 10, (1), whole number.
- BRUNSWIK, E., (1956) *Perception and the representative design of psychological experiments*. Berkeley, California: University of California Press.
- BRUYN, S., (1972) *La perspectiva humana en Sociología*. Buenos Aires: Amorrortu.
- CAMPBELL, D.T., (1979) «Degrees of freedom» ant the case study. In T.D., Cook & C.S., Reichardt (Eds.) *Qualitative and quantitative methods in evaluation research*. Berverly Hills, California: Sage, 49-67.
- CAMPBELL, D.T., & STANLEY, J.C., (1966) *Experimental and quasi experimental designs for research*. Chicago: Rand McNally (trad. cast., 1973).
- CAPARRÓS, A., (1980) *Los paradigmas en Psicología*. Barcelona: Horsori.
- COLL, C., (1981) Algunos problemas planteados por la metodología observacional: Niveles de descripción e instrumentos de validación. *Anuario de Psicología*, 24, (1), 111-131.
- COOK, T.D., & CAMPBELL, D.T., (1976) The design and conduct of quasi-experiments and true experiments in field settings. In M., Dunnette (Ed.) *Handbook of industrial and organizational psychology*, Skokie, Illinois: Rand McNally.
- COOK, T. D., & CAMPBELL, D.T., (1979) *Quasi-Experimentation. Design & Analysis issues for field settings*. Chicago: Rand McNally.
- CONDON, W.S., & OGSTON, W.D., (1967) A segmentation of behavior. *Journal of Psychiatric Research*, 5, 221-235.
- DATAN, N., (1983) Normative or not? Confessions of a fallen epistemologist. In E.J. Callahan & McCluskey (Ed.) *Life-span developmental psychology. Nonnormative life events*. New York: Academic Press, 35-43.
- DENZIN, N.K., (1970) *The research act*. Chicago: Aldine.
- ERICKSON, F., (1977) Some approaches to inquiry in school-community ethnography. *Anthropology and Education Quarterly*, 8, 58-69.
- FILSTEAD, W.J., (Ed.) (1970) *Qualitative methodology*. Chicago: Markham Publications.
- GLASER, B., & STRAUSS, A., (1968) *The discovery of grounded theory: Strategies for qualitative research*. London: Weidenfeld & Nicolson.
- GUBA, E.G., (1978) *Toward a methodology of naturalistic inquiry in educational evaluation*. CSE Monograph Series in Evaluation Number 8. Los Angeles: Center for the Study of Evaluation, University of California.

- GUBA, E.G., & LINCOLN, Y.S., (1982) *Effective evaluation*. San Francisco: Jossey-Bass.
- HABERMAN, S.J., (1978) *Analysis of qualitative data. Vol. 1: Introductory topics*. New York: Academic Press.
- HABERMAN, S.J., (1979) *Analysis of qualitative data. Vol. 2: New developments*. New York: Academic Press.
- HALFPENNY, P., (1979) The analysis of qualitative data. *Sociological Review*, 27, (4), 15-30.
- HUBERMAN, A.M., (1981) Splendeurs, misères et promesses de la recherche qualitative. *Education et Recherche*, 3, 233-249.
- HUBERMAN, A.M., & MILES, M.B., (1983) Drawing valid meaning from qualitative data: Some techniques of data reduction and display. *Quality & Quantity*, 17, 281-339.
- HURSCH, C.J.; HAMMOND, K.R., & HURSCH, J.L., (1964) Some methodological considerations in multiple-cue probability studies. *Psychological Review*, 71, (1) 42-60.
- IANNI, F.A.J., & ORR, M.T., (1979) Toward a rapprochement of quantitative and qualitative methodologies. In T.D., Cook & C.S., Reichardt (Eds.) *Qualitative and quantitative methods in evaluation research*. Beverly Hills, California: Sage, 87-97.
- KNORR-CETINA, K.D., (1981) *The manufacture of knowledge. An essay on the constructivist and contextual nature of science*. Oxford: Pergamon Press.
- KRUGLANSKI, A.W., & KROY, M., (1976) Outcome validity in experimental research: A re-conceptualization. *Representative Research in Social Psychology*, 7, 166-178.
- KVALE, S., (1983) The quantification of knowledge in Education: On resistance toward qualitative evaluation and research. In B., Bain (Ed.) *The sociogenesis of language and human conduct*. New York: Plenum, 433-447.
- KUHN, T., (1962) *The structure of scientific revolutions*. Chicago: Phoenix (trad. cast., 1979).
- LAZARSFELD, P.F., (1955) Some functions of qualitative analysis in social research. *Sociologica*, 1, 355-356.
- LOFLAND, J., (1971) *Analyzing social settings*. Belmont, California: Wadsworth.
- MADGE, J., (1963) *The origins of scientific sociology*. London: Tavistock Publications.
- MASTERMAN, M., (1970) The nature of a paradigm. In I. Lakatos & M. Musgrave (Eds.) *Criticism and growth of knowledge*. Cambridge: Cambridge University Press (trad. cast., 1975).
- MAYOR, J., y SÁINS, J., (1984) Categorización y conceptualización. En J., Mayor (Ed.) *Psicología del pensamiento y del lenguaje*. Madrid: U.N.E.D.
- MILLER, S.I., (1982) Quality & Quantity: Another view of analytic induction as a research technique. *Quality & Quantity*, 16, 281-295.
- MILES, M.B., (1979) Qualitative data as an alternative nuisance: The problem of analysis: *Administrative Science Quarterly*, 24, 590-601.

- MILES, M.B., & HUBERMAN, A.M., (1984) *Qualitative data analysis. A sourcebook of new methods.* Beverly Hills, California: Sage.
- MORGAN, G., (Ed.) (1983) *Beyond method. Strategies for social research.* Beverly Hills, California: Sage.
- MUSSEN, P.H.; CONGER, J.J., & KAGAN, J., (1979) *Child development and personality.* New York: Harper & Row.
- PARLETT, M. & H MILTON, D., (1976) Evaluation as illumination: A new approach to the study of inovatory programs. In G.V., Glass (Ed.) *Evaluation Studies Review Annual*, vol. I. Berverly Hills, California: Sage.
- PATTON, M. Q., (1978) *Utilization-focused evaluation.* Beverly Hills, California: Sage.
- PATTON, M. Q. (1980) *Qualitative evaluation methods.* Beverly Hills, California: Sage.
- PATTON, M. Q., (1984) Data collection: Options, strategies, and cautions. In L., Rutman (Ed.) *Evaluation Research Methods: A basic guide.* Beverly Hills, California: Sage, 39-63.
- REESE, H.W., & SMYER, M.A., (1983) The dimensionalization of life events. In E. J., Callahn & K. A., McCluskey (Eds.) *Life-span developmental psychology: Nonnormative live events.* New York: Academic Press, 1-33.
- REICHARDT, C. S., & COOK, T. D., (1979) Beyond qualitative *versus* quantitative methods, In T.D., Cook & C.S., Reichardt (Eds.) *Qualitative and quantitative methods in evaluation research.* Beverly Hills, California: Sage, 7-32.
- RIECKEN, H.W., & BORUCH, R.F., (Eds.) (1974) *Social experimentation. A method for planning and evaluating social intervention.* New York: Academic Press.
- ROBINSON, W.S., (1951) The logical structure of analytic induction. *American Sociological Review*, 16, 812-818.
- SÁINZ, J., (1984) Hacia un nuevo modelo de categorización. En J., Mayor (Ed.) *Actividad humana y procesos cognitivos.* Madrid: Alhambra, 81-103.
- SMITH, E.E., & MEDIN, D.L., (1981) *Categories and concepts.* Cambridge, Mass.: Harvard University Press.
- SMITH, J.K., (1983) Quantitative versus interpretive: The problem of conducting social inquiry. In E.R., House (Ed.) *Philosophy of evaluation. Philosophy of evaluation.* San Francisco: Jossey-Bass, 27-489.
- SNOW, R.E., (1968) Brunswikian approaches to research on teaching. *American Educational Research Journal*, 5, (4), 475-489.
- SNOW, R.E., (1979) Diseños representativos y cuasi-representativos para la investigación en la enseñanza. En F., Alvira; M.D., Avia; R. Calvo y J.F., Morales. *Los dos métodos de las Ciencias Sociales.* Madrid: Centro de Investigaciones Sociológicas, 189-226 (edic. orig., 1974).
- TURNER, B.A., (1981) Some practical aspects of qualitative data analysis: One way of organising the cognitive processes associated with the generation of grounded theory. *Quality & Quantity*, 15, 225-247.

- VON WRIGHT, G.H., (1979) Two traditions. In J., Bynner & K.M., Stribley (Eds.) *Social research: Principles and procedures*. Bungay, Suffolk: The Open University, 11-16.
- WACHTMAN, E.L., (1982) Storytelling. The narrative structure of evaluation. In N.L., Smith (Ed.) *Communication strategies in evaluation. Vol. 3: New perspectives in evaluation*. Beverly Hills, California: Sage, 89-119.
- WEBB, E.J., (1978) Unconventionality, triangulation, and inference. In N.K., Denzin (Ed.) *Sociological methods. A sourcebook*. New York: McGraw-Hill, 322-328.
- WEIMER, W.B., (1979) *Notes on the methodology of scientific research*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- WEISS, R.S., & REIN, M., (1972) The evaluation of broad-aim programs: Difficulties in experimental design and an alternative. In C.H., Weis (Ed.) *Evaluation Action Programs: Readings in social action and education*. Boston: Allyn & Bacon.
- WOODRUM, E., (1984) «Mainstreaming» content analysis in Social Science: Methodological advantages, obstacles, and solutions. *Social Science Research*, 13, 1-19.

## Incidencia en la práctica educativa de la investigación en el rendimiento

por

Margarita Bartolomé

Después de analizar estos días el estado educativo actual de las investigaciones sobre el rendimiento y las tendencias que se observan desde el punto de vista metodológico, parece natural que nos preguntemos: Y todo ello ¿Qué incidencia tiene en la práctica escolar?. O expresada la cuestión en términos más amplios: ¿Qué *incidencia social* presentan estas investigaciones?

Para quienes son profesores e investigadores a un tiempo, la pregunta emerge de manera espontánea. Pero no podemos olvidar la vertiente *social* de toda investigación científica, y en especial de aquella que se centra en problemas humanos.

En nuestro caso se añade la existencia de factores que hacen más urgentes y necesarias las respuestas a estos interrogantes. La dependencia científica con respecto a la corriente anglo-sajona (constatada fácilmente al observar al bibliografía citada en trabajos de investigación o en reflexiones teóricas) puede aumentar más *la distancia* entre nuestros planteamientos actuales y la práctica educativa basada en una realidad social y escolar muy distinta a la existente en Estados Unidos o Inglaterra. Por citar dos países cuya incidencia, no sólo en el terreno de las Ciencias de la Educación sino incluso en algunos aspectos de la política educativa española parece evidente.

Reconocer este hecho, sin embargo, no significa renunciar a él a priori. Si otros países dedican y han dedicado durante años más medios y recursos a la investigación científica en educación, si cuentan actualmente con una metodología elaborada o unas técnicas interesantes, es lógico que aprovechemos este material que, de algún modo es ya patrimonio de la comunidad humana.

El problema es más bien de adecuación a la *realidad* y de alerta a las características *sociales* sobre las que descansan tales reflexiones científicas y a las que sirven técnicas y metodologías de investigación educativa. ¿Cómo olvidar, por ejemplo, que la evaluación educativa en EE.UU. ha surgido en estos últimos años como reacción a una sofisticación excesiva de las técnicas de medición utilizados en las evaluaciones escolares?... ¿Y que un representante tan calificado

de esta corriente cualitativa, como Guba, ha sido precisamente un gran matemático y conocedor a fondo de los métodos cuantitativos?. Y otro tanto podríamos decir de autores, tan repetidamente citados en nuestras comunicaciones y ponencias como Campbell o Cook.

Estar atentos a la *realidad*, observar los *límites* que imponen nuestras metodologías empleadas, descubrir la *fuente* de nuestros propios errores, *inventar* formas nuevas de profundizar en el conocimiento científico, no sujetarse a *esquemas* preconcebidos, son actitudes propias del científico que lo es realmente.

De ahí que, al hacer balance de las investigaciones realizadas *en nuestro país* quizá sea conveniente que nos acerquemos a su realidad educativa, tal y como funciona hoy, para interrogarla e interrogarnos sobre la educación de nuestros métodos y enfoques.

El segundo factor que, a mi parecer, nos exige esta revisión es el hecho de encontrarnos inmersos, como otros muchos países, en una reforma educativa de amplios alcances. Ahora bien, entre los objetivos que Maravall J.M. (1984, pg. 6) asigna a la reforma se encuentra el de «promover», en fin, una reforma de métodos, contenidos, de programas y pedagogías, capaz de mejorar el resultado de la población escolar, de elevar la calidad de la oferta educativa y de *optimizar el rendimiento en la escuela*. No hay duda que el análisis de los factores que están influyendo más fuertemente en el rendimiento de los alumnos de nuestros centros escolares, en los diversos niveles educativos, podría arrojar una luz decisiva de cara a esa optimización buscada.

Conviene recordar que el proceso por el que los resultados de la investigación actúan de forma significativa en el cambio de conceptos y prácticas educativas suele ser largo y no de efectos inmediatos. Quizá hay que excluir a los docentes que intervienen y participan en la misma investigación. Pero sobre este punto me referiré más adelante.

VARIABLES que mediatizan el proceso son entre otras, el nivel de investigación realizada (básica, aplicada), el modo de divulgación (lenguaje y cauces de publicación utilizados) su adecuación a las *necesidades* sentidas por los que van a utilizar los datos, la forma como estos resultados son conocidos y utilizados en la formación Inicial y Permanente del Profesorado y en los gabinetes donde se gesta la Política Educativa.

Sobre la influencia real de las investigaciones en la Política Educativa de un país, Henriquez, S. (1985, pág. 48) afirma: «nos merece muchas dudas el que los resultados de las investigaciones en educación tengan alguna influencia en la formulación de la política educacional de un país».

Las grandes reformas mundiales en educación que caracterizan los años 60 habían intentado incorporar los planteamientos más innovadores de las Ciencias Pedagógicas en declaraciones de principios que informaban la política a seguir.

El debate –a nivel UNESCO– y la reflexión sobre la educación se habían concentrado en 1971 en el informe «Edgar Faure» de la comisión internacional sobre el desarrollo de la Educación, conocido por el título «Aprender a ser». Pero esta «innovación desde arriba» contrasta con la pobreza de resultados.

(UNESCO 1981). Ciertamente existen cambios significativos en la ampliación y prolongación de la escolaridad, en los contenidos de formación, en la definición a grandes líneas de la educación permanente (aunque a veces se yuxtaponen, sin demasiada coherencia, diversos sistemas de educación)... Pero también aparece un *continuismo en la práctica educativa que transforma y modera extraordinariamente los propósitos más audaces*.

Possiblemente fallaron los elementos claves que mediatisaban el proceso. Las reformas que se perfilan en torno a los 90 han comenzado por introducir correctivos importantes al proceso de innovación.

Para los que participamos de la función investigadora y docente a un tiempo (y «sufrimos» por ello los cambios introducidos por cualquier reforma educativa), el conocimiento de las tendencias que se descubren a nivel mundial puede ayudar a enfocar más realistamente nuestra propia actividad. De los aspectos señalados en los informes prospectivos de la UNESCO (1981) (1984) y de los que se han llevado a cabo en algunos países destacamos dos que tienen incidencia clara en el análisis de las áreas que hemos señalado para nuestro estudio.

1) La búsqueda y propuesta de finalidades nuevas en la educación y que repercutirá indudablemente en la concepción del rendimiento educativo, se lleva a cabo a partir de un análisis crítico de los problemas reales con los que se enfrenta el hombre concreto, en contextos determinados. Lo que se espera de las instituciones escolares es que generen procesos que permitan al niño, al adolescente y al adulto enfrentarse con cuestiones tales como la lucha por la supervivencia física, la grave crisis económica y social, el reconocimiento de la identidad cultural de su propio pueblo; la multiplicidad de conocimientos y la informatización crecientes que requieren el dominio de lenguajes nuevos y la organización selectiva de su propia actividad de estudio; la libertad de criterio frente al progresivo poder de la información, la integridad moral y la apertura ante valores nuevos.

Por otra parte, algunos informes inciden en la necesidad de una mayor exigencia de nivel, que garantice la calidad educativa sin traicionar al tiempo la igualdad de oportunidades. Así en Gardner, D. (1983 pág. 19 se lee: «Recomendamos a los colegios «Colleges» y Universidades que adopten niveles medibles y rigurosos y que se pida más con relación al aprovechamiento académico y a la conducta del alumno y que los Colleges y Universidades eleven sus requisitos de admisión. Esto ayudará a los alumnos a que aprovechen al máximo su educación con materiales que les espoleen y en un ambiente que fomente el aprendizaje y aprovechamiento auténtico».

Este nivel se concreta y especifica a veces excesivamente, intentando «asegurar» nuevas bases que permitan una fundamentación sólida del proceso de aprendizaje permanente. Esta doble concepción de los resultados que se esperan alcanzar (amplia y multidimensional o restringida más a un contenido básico) coexiste en planteamientos actuales.

2) A nivel metodológico se insiste en la reforma desde «abajo», desde los procesos innovadores que se generan en las instituciones escolares y que se lle-

van a cabo por grupos y comunidades educativas a partir de reconocimiento de sus propios problemas. Se acentúa, desde ese enfoque, la importancia de la investigación-acción. También desde amplios programas de investigación básica y aplicada –como el llevado a cabo por la Universidad de Chelsea (Inglaterra)– se aportan datos sobre la conveniencia de adecuar la naturaleza del contenido de las materias científicas a enseñar al diferente tipo y nivel de desarrollo cognoscitivo de los alumnos de cara a una optimización de los resultados de aprendizaje. A modo de ejemplo significativo recordamos que el bajo rendimiento obtenido con el programa de la fundación Nuffield-O nivel, (constituído por estimulantes materiales del área científica) en las «comprehensive schools» pudo explicarse por el carácter selectivo (con respecto al nivel cultural de los sujetos) de las «public schools» donde había sido probado previamente el programa. Un sencillo problema de muestreo había dificultado un ambicioso proyecto de innovación. (Shayer. M. y Adey, P. 1984).

Vamos brevemente, siguiendo el programa de este seminario, a señalar tres áreas que parecen destacarse para analizar al incidencia en la práctica educativa de las investigaciones en el rendimiento.

- 1– En la misma concepción y enfoque de lo que se entiende por rendimiento escolar, tal y como se refleja en la normativa oficial y en la práctica educativa.
- 2– La expresión de esta concepción en la evaluación del rendimiento, observando áreas que se evalúan y modo de llevarlo a término. Siempre en su doble vertiente oficial y práctica.
- 3– En el reconocimiento de factores que afectan al rendimiento, y que pueden o no tenerse en cuenta en la práctica educativa y en la organización de la forma.

No es el objetivo de esta mesa redonda llegar a un análisis exhaustivo de cada una de las áreas sino sugerir reflexiones y abrir pistas y caminos para estudios posteriores de análisis de la realidad.

## **INCIDENCIA DE LA INVESTIGACIÓN SOBRE EL RENDIMIENTO EN LA CONCEPCIÓN Y ENFOQUE DEL MISMO**

En ponencias anteriores (Rodríguez Espinar, S. 1985) se nos ha ofrecido una síntesis de la multidimensionalidad del concepto del rendimiento que se deriva a su vez de la multidimensionalidad de los objetivos de la instrucción. En este aspecto las orientaciones oficiales parecen recoger en los programas básicos de E.G.B, por ejemplo, a través de los objetivos globales para las diversas materias y las orientaciones pedagógicas, una pluralidad y riqueza de resultados esperados que abarcan habilidades cognitivas y no cognitivas.

La práctica escolar indica, sin embargo, que esta multidimensionalidad que puede detectarse ampliamente en preescolar y primer ciclo, va reduciéndose a medida que se avanza la escolaridad. Es decir, las expectativas de los profesores y padres respecto al producto del proceso enseñanza-aprendizaje, va decantándose poco a poco hacia habilidades cognitivas y –dentro de ellas– hacia el dominio de conocimientos culturales y saberes instrumentales, más que hacia la formación de un pensamiento crítico, por ejemplo. En terminología de Botkin, J. (1979) las expectativas esperadas parecen ajustarse más a los resultados de un aprendizaje de «mantenimiento» que «innovador».

Este cambio necesario en la mentalidad del profesor para que los objetivos previstos teóricamente lo sean en la realidad o la distancia existente entre ambos expectativas se acorte, podría constituir una base de investigaciones sobre el rendimiento realmente importante.

## EXPRESIÓN DE ESA CONCEPCIÓN EN LA EVALUACIÓN DEL RENDIMIENTO.

Cuestiones a estudiar en este apartado serían:

¿Dónde se sitúa la evaluación del rendimiento? ¿Qué áreas abarca? ¿Qué criterio se utiliza?.

En segundo lugar deberíamos preguntarnos sobre las *formas* propuestas de evaluar el rendimiento desde las instancias oficiales; qué fundamentación científica presentan; cómo se cubren las necesidades técnicas instrumentales desde la investigación actual; qué elementos deberían ser potenciados; qué formación inicial y permanente del profesorado se necesitaría para «una valoración *objetiva y continua* del rendimiento». (Circular del Director General de Educación Básica. Madrid 26 julio de 1985).

En relación al primer tipo de cuestiones parece clara la distinción en algunas Orientaciones oficiales entre la valoración global de la madurez personal del sujeto (interés por el trabajo, seguridad en sí mismo, relación positiva con los compañeros y adultos), la adquisición de hábitos y técnicas de trabajo y el *rendimiento objetivo* de los aprendizajes básicos. (Véase Direcció General d'Ensenyament, 1983).

Cuando analizamos las orientaciones ofrecidas a los profesores para la evaluación de las tres áreas, desde el punto de vista del contenido, encontramos, sin embargo, un cierto esfuerzo por integrar aspectos cognitivos y no cognitivos al describir los criterios de rendimientos referente a aprendizajes básicos al menos en los dos primeros ciclos de E.G.B. (véase Direcció General d'Ensenyament, 1982).

Hay que reconocer que sobre este material y la manera de utilizarlo los profesores, las dificultades halladas, las operativizaciones no comprendidas o no establecidas con rigor, el paso de *criterios cualitativos* de rendimiento multidimensionales al criterio de *promoción* o a las *calificaciones finales*, existe un vacío

considerable en nuestra investigación educativa. En todo este proceso, la opinión y vivencia del profesorado no ha sido recogida suficientemente ni aprovechada.

Las investigaciones tienden a utilizar el criterio de rendimiento basado en calificaciones o en promoción escolar, cuando se refieren al rendimiento en general. Algunas, en el campo lingüístico sobre todo, se basan en pruebas elaboradas, que no suele ser de uso común entre los educadores.

Este último punto nos introduce en un segundo tipo de cuestiones sobre la forma de obtener indicadores válidos del rendimiento.

Se aprecia una tendencia *significativa* a insistir en metodologías *cualitativas* al señalar los instrumentos que conviene utilizar para la evaluación formativa del rendimiento, desde instancias oficiales. Así en la última circular de la Dirección General de E. Básica (26.7.85), al tratar este punto se afirma la necesidad de utilizar «*instrumentos de medición adecuados*, tales como: registros de observación sistemática, listas de control, escalas de valoración, anecdotarios, fichas de seguimientos de distintos aprendizajes y áreas, trabajos personales o de equipo... En consonancia con el párrafo anterior la aplicación de pruebas o exámenes, destinadas a comprobar de modo permanente el dominio de los programas no resulta conveniente y su utilización esporádica será objeto de un tratamiento similar al señalado para otros instrumentos». El párrafo anterior evidencia el estado de opinión que va generalizándose en nuestro país, favorable a las técnicas cualitativas, así como indentificaciones que podrían ser fácilmente deslindables desde una reflexión seria sobre la búsqueda de indicadores válidos y fiables del rendimiento.

Apenas hemos cruzado el umbral de una *mínima* tecnificación a la hora de utilizar los educadores algún tipo de instrumentos cuantitativos elaborados por ellos mismos, las pruebas objetivas, por ejemplo. Esta realidad se evidencia en la investigación de De la Orden Hoz. A (1981) sobre las características de las pruebas de 8.<sup>o</sup> curso de Educación General Básica que, por prescripción legislativa, deben elaborar los centros. En ella se constató, entre otros hallazgos que más del 60% de los items de la subprueba de lengua española, estaban dirigidos a evaluar objetos clasificables en la primera categoría –conocimiento– de la taxonomía de Bloom.

Junto al rechazo de instrumentos de escasa validez y objetividad, se desechan otros que podrían ayudar poderosamente a los profesores, como los tests referidos al criterio, para conocer a través de la evaluación formativa el modo paulatino de dominio de determinadas dimensiones del rendimiento. En este seminario, Anguera. M.T. (1985) nos ha ofrecido una valoración ponderada de las posibilidades de las metodologías cualitativa y cuantitativa aplicadas al rendimiento, que arrojan una luz suficiente sobre cuestiones tales como la *objetividad* en ambas metodologías, las *descripciones* ofrecidas por los datos cualitativos, que deberían conjugarse con *mediciones* cuantitativas, y las *condiciones* para que ambos vías de conocimiento sean válidas y fiables.

El problema que aquí subyace –como en los apartados anteriores– es el de

la formación *inicial y permanente* del profesorado para introducirlo en el conocimiento crítico de los instrumentos y en la familiaridad de sus usos.

La inexistencia de personal especializado técnico en Centros educativos agrava además el problema, ya de por sí preocupante, de una formación deficiente. En esta situación resulta más fácil rechazar las técnicas afirmando que no se cree en ellas que confesar la ignorancia e incapacidad de utilizarlas.

Cuatro son, a mi parecer, las responsabilidades que nos compete asumir a los investigadores en este tema:

- a) Trabajar intensamente por construir instrumentos bien sean pruebas criteriales objetivas o técnicas de observación de acuerdo con las orientaciones y objetivos que van apareciendo desde instancias oficiales. En este sentido el área del lenguaje parece ofrecer ya realizaciones prometedoras.
- b) Procurar, en la elaboración de instrumentos técnicos, que no exista una reducción de la multidimensionalidad del concepto de rendimiento. La elección de unas técnicas u otras puede venir urgida por la naturaleza misma del área a evaluar.
- c) Involucrar, en esta tarea a profesores en ejercicio que pueden ofrecer su experiencia y práctica a la hora de seleccionar indicadores. Colaborar asimismo con otros organismos en la tarea de informar y formar al profesorado en las posibilidades y uso de esas técnicas.

En esta difusión convendría evitar concepciones que han podido tener su origen en un uso inmoderado y erróneo de la tecnología. Por ejemplo, el recelo hacia una posible deshumanización de la función educadora por el uso de tests (tan frecuente en ciertos ámbitos) o, al contrario, la creencia de que las técnicas de observación son muy «fáciles» de utilizar y cualquiera puede construirlas y valorarlas. En este sentido cobra importancia decisiva el *lenguaje* y la *práctica* que acompaña dicha información y/o formación del profesorado.

- d) Detectar las características diferenciales de grupos y personas que pueden afectar al tipo y forma de técnicas instrumentales que conviene emplear en la medición del rendimiento o descripción del mismo.

## INCIDENCIAS DE LAS INVESTIGACIÓN SOBRE FACTORES QUE AFECTAN AL RENDIMIENTO.

Si la investigación sobre construcción y uso de instrumentos de evaluación puede afectar a la larga a la misma concepción del rendimiento escolar que subyace a dichos instrumentos, las investigaciones sobre factores, que afectan al rendimiento y la elección de modelos determinados, pueden cuestionar desde los

enfoques educativos promovidos por las reformas hasta la misma eficacia de la educación como factor nivelador de las desigualdades sociales. Baste, por ejemplo, recordar el uso que se hacía en las alternativas pedagógicas en el 75 de los factores sociológicos frente a los psicológicos como determinantes de las de las diferencias halladas en nivel y calidad de logros escolares.

No hay duda de que el principio de igualdad de *oportunidades*, que está a la base de numerosas políticas educativas evoluciona progresivamente desde un enfoque liberal a una concepción que brota de la constatación de una desigualdad existente a pesar de las reformas en el campo de la educación. Para que exista igualdad de oportunidades, los sujetos deben recibir trato desigual. Y ello se reconoce, desde los movimientos culturalistas que, al reivindicar la consideración en el ámbito educativo de la existencia de un criterio igualitario de rendimiento; desde los planteamientos de educación compensatoria y desde enfoques que parten de la profundización de los diferentes estadios del desarrollo cognoscitivo de los sujetos para estimarlos como indicadores del éxito en consecución de objetivos apropiados a cada nivel y cuya exigencia ha sido analizada.

Las tendencias educativas, que parecen dibujarse a nivel mundial sobre este reconocimiento son:

1- El principio de equidad se vincula al de máxima calidad de la educación, lo que significa intentar que todas las personas desarrollen su *potencial máximo*, alcanzado así un nivel de logro adecuado a dicho potencial. Se superan por tanto los modelos estrictamente ambientalistas de Bloom, B.S. (1964). Dentro de un mismo contexto social el modelo psicológico adquiere mayor relieve. El problema del *diagnóstico precoz*, se acentúa en la práctica escolar, ya que éste adquiere una función claramente predictora del rendimiento. Ahora bien ¿Con qué instrumentos cuentan los educadores para realizarlo? ¿Qué ayudas técnicas está proporcionando la investigación sobre los factores determinantes del rendimiento, para una selección adecuada de las variables?. Algunos problemas, siguen sin acabar de resolverse. Por ejemplo la cuestión del *umbral mínimo* de una variable, por debajo del cual su incidencia en el rendimiento es considerable, pero cuya superación vuelve no significativa la relación entre esa variable y el criterio. En segundo lugar, la búsqueda de indicadores múltiples para algunas variables latentes y la existencia de interacciones múltiples que hace extremadamente compleja la reducción de la realidad a modelos. En este seminario se nos ha ofrecido algunas vías de solución que deberán contrastarse una y otra vez mediante réplicas de investigación adecuadas.

2- Se reconocen, dentro de las variables de proceso que producen una diferenciación individual en el aprendizaje, no sólo las que provienen de los sujetos sino las que proceden de la actuación del profesor, del tipo de recursos utilizado (introducción o no de los nuevos lenguajes...) y de las variables organizativas. El enfoque ecológico y la consideración del *clima de clase* como un factor fundamental que afecta al rendimiento, está teniendo cada vez mayor relevancia tanto desde un enfoque psicométrico (Villar Angulo L.M y Marcelo. C. 1985) como etnográfico.

3- Los programas de educación compensatoria para grupos desfavorecidos (medio rural, suburbano...) tienen una doble fundamentación: de una parte, la conexión estrecha educación-sociedad obliga a plantear la resolución de problemas en contexto cada vez más amplios para que sean resolubles. Las variables no escolares han de introducirse en el modelo de intervención; de otra, se reconoce la efectividad de variables educativas específicas, sobre todo aquellas que hacen referencia a modelos cognitivos, conociendo más profundamente los factores asociados a su proceso y evolución. En España detectamos una falta grave de investigaciones en ese terreno. Hay que tener en cuenta el factor *tiempo* al diseñar tales investigaciones, ya que la efectividad y permanencia de los efectos de los programas en el rendimiento de los sujetos puede observarse a través de medidas repetidas en tiempos sucesivos. Además en los modelos de intervención se está realizando un esfuerzo evidente por no asimilar los conceptos de tratamiento diferencial con los de educación segregadora o compartimentalizada. Así, se vincula una mayor profundización en las características peculiares de los sujetos con diversos tipos de deficiencia a los principios de normalización e integración de éstos en el ámbito escolar y social. Desde el punto de vista de la investigación estos enfoques sugieren áreas nuevas de trabajo donde el análisis de casos y los procedimientos clínicos pueden combinarse con métodos más extensos de tipo ex-post-facto, que pueden servir como referencia. También es interesante la combinación de metodologías cualitativas y cuantitativas sugeridas en las obras de Cook. T y Reichard. Ch. S. (1979) o de Denzin N. (1978).

De todos los anteriores planteamientos vuelven a surgir interrogantes: ¿Cómo conseguir que los profesores se involucren más en acciones que supongan una reflexión sobre lo que vienen realizando? ¿Es suficiente el conocimiento de los resultados de la investigación sobre factores determinantes del rendimiento para abordar con más eficacia los problemas planteados entorno a la optimización de aquel?. El porcentaje de fracaso escolar en España se sitúa en un rango bastante superior al de otros países europeos. ¿Podremos disminuirlo desde un trabajo que intente ofrecer soluciones a estos problemas?

Personalmente entiendo que una de las vías más importantes para afrontar con eficacia estas cuestiones estriba de nuevo en la formación del profesorado. Y ello en una doble dirección. Si la investigación sobre los factores que inciden en el rendimiento puede iluminar la práctica educativa en los aspectos anteriormente mencionados, una investigación-acción en la escuela, atenta a descubrir cómo funciona la actividad educativa y de qué forma pueden irse solucionando algunos problemas del proceso enseñanza aprendizaje, arrojaría luz indudablemente para la elaboración previa de un modelo causal que se ajuste más a la realidad. No hay duda de que en el último curso se ha notado una preferencia en torno a este tema en los ambientes pedagógicos. ¿Nos enfrentaremos a él como una moda más –con sus detractores o defensores a ultranza– o seremos capaces de favorecer pacientemente estos procesos de trabajo sobre la realidad educativa, deseando no sólo comprenderla sino mejorarla, (Kemmis. S. y Karr. W. 1983) y aprendiendo unos de otros como llevar a cabo esa transformación?

He aquí un programa de trabajo que puede iniciarse en algunos de nuestros centros universitarios y desarrollarse en otros. A través de las páginas de la *Revista de Investigación Educativa* iremos descubriendo en los próximos años los frutos de esta aportación científica al problema que nos ocupa.

#### REFERENCIAS BIBLIOGRÁFICAS

- BLOOM. B.S. (1964): *Stability and change in human characteristics* Wiley, New York.
- BOTKIN. J.W., ELMANDJRA. M y MALITZA. M. (1979): *Aprender, horizontes sin límites*. Santillana, Madrid.
- COOK. T. and REICHARD. CH. S. (1979): *Qualitative and Quantitative methods in Evaluation Research*. Seage, Beverly Hills, California.
- DE LA ORDEN. A. (1983) «Evaluación de textos escolares.» *Revista de investigación educativa*. 2. V. 1.240-259.
- DENZIN. N. (1978): *Sociological Methods. A soucerbook*. McGraw Hill New York.
- DIRECCIÓN GENERAL DE EDUCACIÓN BÁSICA. (1985): Circular por la que se regula el comienzo del curso 1985-86 en los centros escolares dependientes del Ministerio de Educación y Ciencia. *Escuela española*. 2.781. 12-13.
- DIRECCIO GENERAL D'ENSENYAMENT PRIMARI. (1983): *Orientacions i programes. Cicle mitjà d'educació general bàsica*. Secretaría técnica, Barcelona.  
– L'avaluació al Cicle Mitjà. (1982): *Full d'informació a les escoles*. 22 9-27.
- GARDNER. D. (1983): *A nation at rik: The imperative for educational reform*. Departament of Education, Washington.
- HENRIQUEZ. S. (1984): Tendencias actuales y futuras sobre la formación de profesores en Europa Occidental. *Studia pedagógica*. 14. 39-55.
- KEMMIS.S and CARR. W. (1983): *Becoming critical: Knowing through Action Research*. Deakin, Victoria.
- MARAVALL. J.M. (1984): *La reforma de la enseñanza*. Laia, Barcelona.
- SHAYER. M y ADEY. P. (1984): *La ciencia de enseñar ciencias. Desarrollo cognoscitivo y exigencias del currículo*. Narcea, Madrid.
- SHWARTZ. H. y JACOB. J. (1984): *Sociología cualitativa*. Trillas, México.
- UNESCO. (1981): *Las finalidades de la educación*. UNESCO, Paris.  
– (1981): *Las reformas de la educación. Experiencias y prospectivas*. UNESCO, Paris.
- VILLAR ANGULO. L.M. y MARCELO. C. (1985): «Estudio y evaluación del ambiente escolar en relación a otras variables: contextualización del rendimiento escolar». *Bordón*. 257. 256-273.

## Evaluación con respecto al criterio en matemáticas (Proyecto E.A.O.-TOAM)

por  
*Inmaculada Bordas*

Este artículo presenta un estudio –diseño cuasiexperimental– que se llevó a término en el Proyecto E.A.O. –TOAM, del Departament d'Ensenyament de la Generalitat de Catalunya (Véase artículo «Evaluación del rendimiento en Aritmética» del Dr. Benito Echeverría) en el curso 1984-1985, y cuya finalidad es:

1. Controlar a través de un instrumento externo las diferencias existentes entre el «grupo control» y el «grupo experimental» en cuanto:
  - a.– Proceso global de aprendizaje.
  - b.– Niveles didácticos: «Concepto», «Habilidad mecánica» y «Razonamiento»
  - c.– Aprendizaje de las distintas áreas de contenido de la materia.
2. Contrastar los resultados de las Pruebas TOAM, con los de un instrumento evaluativo externo.

Para la consecución de estos dos objetivos fue necesario el trabajo previo de la confección de tres pruebas –una para cada Ciclo de E.G.B.–, así como su análisis y la comprobación de su fiabilidad y validez.

### LAS PRUEBAS

#### Contenido de las pruebas. Condicionantes.

Las tres pruebas creadas para llevar a término esta investigación fueron confeccionadas atendiendo a:

\* *La materia: Matemáticas*

Esencialmente se tuvo presente la estructura de esta área del saber, su lógica

docente y los niveles didácticos de aprendizaje («Conceptual», «habilidad mecánica» y «razonamiento».)

\* *Los examinados*

La psicología de los alumnos de 2.<sup>º</sup>, 5.<sup>º</sup> y 8.<sup>º</sup> de E.G.B. en los aspectos cognoscitivos, activos y afectivos; su capacitación matemática, lectora y de expresión escrita.

\* *Los objetivos mínimos obligatorios de cada ciclo del área.*

Fueron determinantes en este ámbito las directrices dadas por el M.E.C.y la Conselleria d'Ensenyament de la Generalitat de Catalunya, así como los textos vigentes en el año en curso y más utilizados en nuestras escuelas.

### Definición de las pruebas. Presentación.

Partiendo de las bases expuestas se determinaron:

\* *Los objetivos de cada una de las pruebas en las distintas subáreas*

Estos objetivos fueron definidos en cada ciclo a dos niveles: «objetivo pedagógicos», y «objetivos de comportamiento» (Bloom)

		C. INICIAL (N.I.=45)				C. MEDIO (N.I.=60)				
		Lógica	Numer.	Medid.	Geome.	Lógica	Numer.	Medid.	Geome.	
Contenido	Nivel Didac.	13 %	20 %	7 %	13 %	53 %	7 %	13 %	10 %	
		—	18 %	—	—	18 %	3 %	27 %	—	
		—	24 %	5 %	—	29 %	5 %	13 %	10 %	
		13 %	62 %	12 %	13 %	100 %	15 %	53 %	20 %	
		C. SUPERIOR (N.I.=60)								
Contenido	Nivel Didac.	Numer.	Func.	Polif.	Propor.	Medid.	Geom.	Estad.		
		18 %	5 %	3 %	2 %	3 %	10 %	2 %	43 %	
		12 %	—	—	—	—	—	—	12 %	
		15 %	6 %	2 %	5 %	—	17 %	—	45 %	
		45 %	11 %	5 %	7 %	3 %	27 %	2 %	100 %	

TABLA I  
Distribución porcentual de los ítems en cada Prueba

*\* La magnitud de las pruebas*

A fin de determinar el número de ítems se buscó el equilibrio entre los objetivos de cada prueba, «la fatiga discente» en la realización de la misma y el rendimiento. La TABLA I refleja estos valores, así como la distribución porcentual de los mismos en las distintas áreas y niveles didácticos.

*\* Los ítems.*

Cada ítem se presentó al discente como unidad independiente y diferenciada, con un lenguaje propio del nivel-área al tiempo que familiar. Las construcciones gramaticales eran simples y sencillas.

A su vez los reactivos tenían diferentes niveles de dificultad debido principalmente a los factores de «complejidad de aprendizaje», «tiempo-práctica» y «base o no de ulteriores aprendizajes».

La respuesta que debía dar el alumno era de distintos tipos («completar», «gráficos», «bien-mal», de «selección múltiple» y «apareamiento») a fin de evitar la reiteración y el aburrimiento.

*\* La forma de las pruebas.*

En los diferentes ciclos, las pruebas se ofrecieron en castellano y catalán (la lengua no debe ser una dificultad en la realización del trabajo). Los ítems y las áreas de contenido se distribuían tal como nos indica la TABLA II.

CICLO	CUADERNO	N.º ITEMS	SUBAREA DE APRENDIZAJE
INICIAL	1	24	-Sistema numerac.: operatividad -Geometría -Lógica
	2	21	-Geometría -Lógica -Sistema de numeración -Medidas
MEDIO	1	34	-Sistema numeración: operativ. -Medidas.
	2	26	-Lógica. -Sistema numeración (base 10/base≠10) -Geometría
SUPERIOR	1	34	-Numeración (Div. N. Dec., Frac.,...) -Polinomios
	2	26	-Num. Z -Funciones. Proporciones -Medida. Geometría -Estadística

TABLA II  
Distribución de contenidos e ítems en los diferentes cuadernos

#### \* *Las Pruebas*

Las tres pruebas quedaron estructuradas en seis cuadernos cuyos componentes o ítems se determinaban esencialmente por la definición de objetivo, la subárea de contenido, el nivel didáctico, el aspecto mental a que atiende, el tipo de solución y el margen de error.

#### **Definición del método de aplicación**

Las directrices a seguir se dirigían al examinando y al examinador, incluyendo en las de éste los resultados-respuestas de los ítems ejemplo.

Cabe destacar el hecho de que las pruebas se efectúan en dos sesiones como mínimo a fin de evitar la fatiga y obtener el máximo rendimiento discente.

#### **La corrección**

El material utilizado para la corrección consiste en unos cuadernos—plantilla en donde, además de la respuesta, se explicita el «margen de error» del ítem. La puntuación para cada ítem es de 0 o 1 según la respuesta sea omitida, errónea o bien acertada. La puntuación global para cada alumno es igual a la suma de aciertos.

### **ANALISIS DE LAS PRUEBAS**

El estudio de las pruebas se realizó en dos fases: la primera corresponde a un trabajo piloto y la segunda se efectuó mediante un análisis de las pruebas rectificadas en una muestra de 1264 alumnos.

#### **Estudio piloto.**

Se llevó a término en la escuela Marqués de la Pobla de Capellades (Barcelona), en los ciclos inicial (C.I.) y medio (C.M.) en un total de 134 alumnos (68 y 66 respectivamente). Los resultados se concretan en:

#### \* *Análisis global de las Pruebas*

- Los índices de fiabilidad hallados por el M. de HOYT eran elevados (C.I.: .79; C.M.: .85) por lo que las pruebas en su globalidad no deberían ser reconsideradas.
- Las pruebas discriminaban a los sujetos y a los elementos al nivel 1%.
- Su validez era significativa puesto que: a) representaban los objetivos de una secuencia determinada; b) relacionadas las puntuaciones globales de

las pruebas y las del profesor mostraban correlaciones significativas entre ambas (.55 y .57 respectivamente).

\* *Análisis de los ítems*

Estudiados en cada ítem los índices de dificultad (I. porcentual) y de discriminación (I. Lindeman) y teniendo presente las preguntas y dificultades de los alumnos en la realización de las Pruebas se consideró que debían rectificarse 15 ítems de C.I. y 17 del C.M. por las causas que se explicitan en la TABLA III.

CICLO	N.º ITEM	CAUSA	CORRECCION
INICIAL	13	El alumno espera datos exactos.	Remodelación texto. Determinar enunciado.
	14-15 31-36	Baja comprensión del enunciado gráfico.	Gráfico
	44	El alumno debe escribir cuatro palabras. Lentitud.	Redacción de nuevo texto a fin de cambiar el tipo de respuesta.
	3-4 8-12 18-22 23-38 39	Dificultades situacionales. Lentitud en la corrección.	Diferenciar «espacios» <i>b</i> y <i>c</i> . Incluir la palabra «respuesta» en el espacio <i>c</i> .
	1-39 54	Baja comprensión del enunciado gráfico.	Gráfico
MEDIO	23	No queda clara la disposición del ítem. Difícil corrección.	Gráfico
	5-6 7-9 14-18 19-26 28-33 34-52 53.	Dificultades situacionales. Lentitud en la corrección.	Diferenciar «espacios» <i>b</i> y <i>c</i> . Incluir la palabra «respuesta» en el espacio <i>c</i> .

TABLA III  
Items que deben ser rectificados después del estudio piloto.

**Estudio empírico de la prueba definida**

\* *La muestra*

Las pruebas rectificadas del C.I. y del C.M. conjuntamente con las del C.S. se administraron a una muestra de 1264 alumnos de nueve centros públicos de

Barcelona y su provincia, según se indica en la TABLA IV. Cada uno de estos centros ha sido descrito por el Dr. B. Echeverría en el artículo citado anteriormente.

Ciclo \ Centro Escolar	JULIA MINGUELL	PERE VIVER	ELS HORTS	BISBE BERENGUER	BERNAT DESCLOT	MILAGROS CONSARNAU	PLANES I CASALS	J. MARQUES CASALS	PARDO BAZAN	Total	Niños	Niñas	Edad media
Ciclo													
INICIAL	—	111	58	61	48	62	—	63	52	455	227	229	7,96
MEDIO	—	102	63	57	56	33	—	65	59	435	237	197	11,22
SUPERIOR	34	86	53	—	27	54	22	53	45	374	205	169	14,02
Total	34	299	174	118	131	149	22	181	156	1264	669	595	—

TABLA IV  
Centros y número de alumnos que constituyen la muestra

\* *Datos y resultados análisis*

Realizadas las pruebas por todos los alumnos, los datos fueron codificados y transcritos en cinta magnética con el fin de ser procesados por ordenador. Esencialmente, los estudios que se efectuaron fueron:

a. Un análisis descriptivo de cada una de las tres muestras cuyos estadísticos quedan reflejados en la TABLA V.

Ciclo	Mínim.	Maxim.	M. Arit.	Med.	Mo.	Desv. St.	Error. St.	Kurtosis	Asimet.	• Percentiles					
										25	50	75	30	60	90
C.I.	6	43	31,1	32	37	7,4	0,3	-,12	-,62	27	32	37	28	34	40
C.M.	9	59	38,7	39	34	10,1	0,4	-,02	-,37	33	39	46	34	41	52
C.S.	1	51	20,9	20	18	9,3	0,4	+,09	+,52	14	20	27	15	22	34

TABLA V  
Resultados del estudio estadístico descriptivo.

b. Obtención de los coeficientes de fiabilidad mediante el método de Crombach, el cual nos dió los valores de: C.I.:  $\alpha = .89$  ( $N=455$ ); C.M:  $\alpha = .91$  ( $N=435$ ); y, C.S.:  $\alpha = .89$  ( $N=374$ ).

Estos valores se ratificaron por el método de Hoyt en el que se obtuvo respectivamente por prueba/ciclo .88, .90 y .89.

c. Estudio de la validez. Recordemos que toda prueba puede considerarse válida siempre que sea instrumento que mida el aprendizaje del alumno en los distintos niveles didácticos en las respectivas áreas de conocimiento, es decir, en cuanto mida en qué grado se han alcanzado los objetivos propuestos. Desglosando este concepto (Lindeman) se evidenció en las pruebas validez de contenido y de construcción porque:

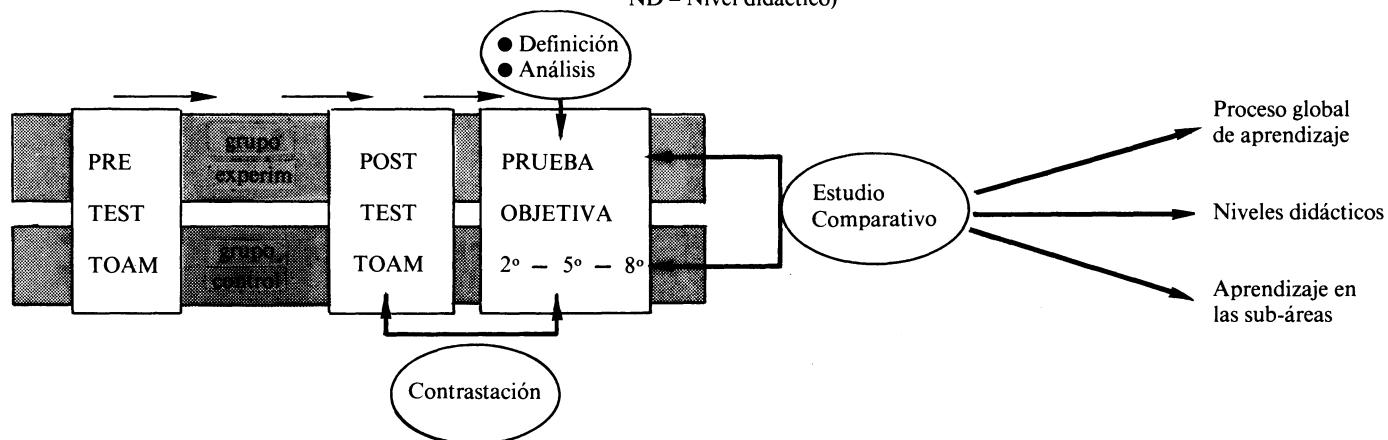
- Representaban fielmente los objetivos de la secuencia didáctica presentada, la cual, como ya se dijo seguía las Orientaciones del M.E.C. y de la Generalitat de Catalunya así como los libros de texto más utilizados en los centros.
- Cada prueba, como unidad e ítem por ítem, fue examinada por 5 jueces que analizaron los contenidos, la secuenciación, el lenguaje y la forma de presentación. Cada sentencia debía ser aprobada como mínimo por 4 de los cinco jueces.
- La correlación entre las pruebas y las puntuaciones del profesor en el estudio piloto fueron significativas.

d. Análisis de cada uno de los ítems. Cada ítem se estudió desde distintos puntos de vista. De ello se obtuvo que había una gran oscilación entre los índices de dificultad (98'7 a 2%) lo cual es lógico por la naturaleza de la materia, su aprendizaje secuencial y las diferencias individuales de aprendizaje (véase TABLA VI). También se dedujo que todos los ítems eran válidos ( $r_{bp}$  entre la puntuación global de la prueba y la de cada ítem) excepto tres (C.I.: ítem n.º 27, área de lógica, nivel «concepto»; C.M.: ítem n.º 36, área de medidas, nivel «razonamiento»; y del C.S.: ítem n.º 34, área Numeros Z-, nivel de «concepto») cuyo índice no es significativo al 5%; no obstante se consideró que estos ítems se revisarían en una nueva fase de estudio y que no afectaban significativamente a los resultados globales. Por último, vimos que la aportación de cada ítem en la fiabilidad global de la prueba era semejante lo cual indicaba que la constitución interna de la misma era muy homogénea.

Como resumen y final de todo este estudio, se concluyó que estas pruebas podrían ser utilizadas como «Instrumento externo de control» para conocer las diferencias y paralelismos del aprendizaje matemático entre los grupos «experimental» y «control» en la investigación Proyecto EAO-TOAM.

		LOGICA	NUMERACION	FUNCION	POLINOM.	PROPOR.	MEDIDA	GEOMET.	ESTAD.
CI	% N. <sup>o</sup> ND	2,6-55,3 27-29 a a	2-67,5 1-21 b b	—	—	—	15,4-50,4 36-38 a a	7,9-35,1 42-43 a a	—
CM	% N. <sup>o</sup> ND	2,5-65,5 42 41 c c	3,9-65,3 7 20 a b	—	—	—	12,9-80,5 24 31 a c	30,6-72 58 54 a a	—
CS	% N. <sup>o</sup> ND	— — —	27,8-97,3 8 23 c a	51,6-95,2 25 31 a c	41,4-98,7 33 34 a c	58,6-88 41 40 c c	27-66,8 42 43 a a	21,7-97 54 59 a c	84,8 60 a

TABLA VI  
 Índices de dificultad. Valores máximos y mínimos  
 (%-Índice de dificultad; N.<sup>o</sup> -Número del ítem al que corresponde;  
 ND – Nivel didáctico)



ESQUEMA  
 Aspectos de estudio de la Prueba Objetiva respecto al esquema general

## ESTUDIO COMPARATIVO ENTRE EL GRUPO EXPERIMENTAL Y EL GRUPO CONTROL

Este apartado presenta el estudio comparativo que se efectuó entre dos grupos –recordemos el Proyecto EAO-TOAM– y cuya finalidad fue la de detectar las diferencias en el aprendizaje matemático en cuanto a «proceso glocal», a «niveles didácticos» (n. conceptual, n. de habilidad mecánica y n. de razonamiento) y a «aprendizajes en las distintas sub-áreas de contenido».

Los instrumentos que se utilizan fueron las tres pruebas cuyo estudio se ha mostrado anteriormente.

El ESQUEMA muestra la ubicación de este trabajo dentro el Proyecto así como los diferentes núcleos de estudio.

### Muestra

La muestra la conformaron dos grupos: un primer grupo o «experimental» constituido por alumnos que habían realizado sus aprendizajes de aritmética con ayuda del programa TOAM; y un segundo grupo o «control», que efectuaba sus clases con los métodos y recursos habituales y que no había utilizado EAO.

Ambos grupos-muestra los analiza más exhaustivamente el Dr. Echeverría en el informe citado, por lo que creemos innecesario repetirlo. No obstante recordemos los datos que se exponen en la TABLA VII.

	CICLO INICIAL		CICLO MEDIO		CICLO SUPERIOR	
	Gr. EXP.	Gr. CONT.	Gr. EXP.	Gr. CONT.	Gr. EXP.	Gr. CONT.
N.º Alumnos	284	172	283	152	204	170
Niños	139	88	139	98	119	86
Niñas	145	84	144	54	85	84
Edad Media (Junio, 1985)	8,07	8,78	11,33	11,02	14,08	13,95

TABLA VII  
Número de alumnos en los diversos grupos de la muestra

### Prueba: estudio estadístico comparativo entre los grupos

#### \* La «Puntuación global».

En cada grupo esencialmente se obtuvieron los estadísticos que se presentan en la TABLA VIII, y de los que se dedujo que:

- En el aprendizaje de matemáticas se presentaban mejores puntuaciones en el grupo experimental que en el control.
- Donde había una mayor diferenciación entre grupos era en el Ciclo Ini-

cial (puntuación máxima teórica: 44), y en el Medio (puntuación máxima teórica: 60). En el Ciclo Superior era inferior (véase  $\bar{X}$ ,  $Me$  y  $Mo$ .).

- No había diferencias significativas entre los valores máximos y mínimos en los dos grupos de cada Ciclo.
- La dispersión respecto al valor medio era más reducida en los grupos experimentales.

CICLO	INICIAL		MEDIO		SUPERIOR	
GRUPO	Exp.	Cont.	Exp.	Cont.	Exp.	Cont.
Punt. Media	32,57	28,87	41,44	33,70	22,98	18,50
Me	34	29	41	34	22	17
Mo.	38	27	48	34	20	16
Kurtosis	+0,55	-0,33	-0,11	-0,12	+0,28	+0,19
Asimetría	-0,94	-0,23	-0,32	-0,30	+0,49	+0,70
Des. stand.	7,21	7,29	9,13	10,08	9,06	9,19
L <sub>i</sub>	6	9	10	9	2	1
L <sub>s</sub>	43	43	59	56	51	48

TABLA VIII  
La «puntuación global». Análisis estadístico descriptivo.

Con el fin de detectar con una mayor exactitud la significación de estas diferencias se procedió a un estudio de comparación de medias (TABLA IX). El resultado respecto a cada uno de los Ciclos fue significativo con un grado de .000.

	N. <sup>a</sup> Ind.	$\bar{X}$	Desv. Stn.	Err. Std.	T	GL	Prob.
C. I.	G. Exp.	284	32,57	7,21	0,42	5,22	.000
	C. Cont.	172	28,87	7,24	0,55		
C. M.	G. Exp.	283	41,44	9,13	0,54	8,12	.000
	G. Cont.	152	33,70	10,08	0,81		
C. S.	G. Exp.	204	22,98	9,06	0,63	4,73	.000
	G. Cont.	170	18,50	9,19	0,70		

TABLA IX  
La «puntuación global». Estudio de comparación de medias entre los grupos experimental y control

\* *Estudio de las puntuaciones en los diversos «niveles didácticos» de aprendizaje.*

Con el fin de profundizar más en este estudio comparativo se consideró importante el detectar si había diferencias significativas entre los grupos respecto a los tres grandes niveles didácticos de aprendizaje matemático, es decir, en el nivel «conceptual», en el de «habilidad mecánica» y en el de «razonamiento». Los resultados –contrastación de medias– son los que se indican en la TABLA X.

CICLO	NIVEL DIDACTICO	GRUPO	N.º IND.	$\bar{X}$	DESV. STAN.	ERR. STAN.	T	G.L.	PROB.
C.I.	«Concept.»	Exp.	284	18,01	3,75	0,22	3,79	454	.000
		Cont.	172	16,59	4,03	0,30			
	«Hab. Mec.»	Exp.	284	5,99	1,67	0,10	6,71	454	.000
		Cont.	172	4,93	1,56	0,11			
	«Razonam.»	Exp.	284	8,56	2,64	0,15	4,51	454	.000
		Cont.	172	7,41	2,66	0,20			
C.M.	«Concept.»	Exp.	283	17,43	4,09	0,24	7,70	433	.000
		Cont.	152	14,30	3,93	0,31			
	«Hab. Mec.»	Exp.	283	13,16	3,17	0,18	6,92	433	.000
		Cont.	152	10,76	3,92	0,31			
	«Razonam.»	Exp.	283	9,78	2,91	0,17	6,88	433	.000
		Cont.	152	7,6	3,21	0,26			
C.S.	«Concept.»	Exp.	204	11,16	4,26	0,29	3,86	372	.000
		Cont.	170	9,38	4,61	0,35			
	«Hab. Mec.»	Exp.	204	3,68	1,93	0,13	3,50	372	.001
		Cont.	170	2,98	1,89	0,14			
	«Razonam.»	Exp.	204	8,22	4,16	0,29	4,99	372	.000
		Cont.	170	6,16	3,70	0,28			

TABLA X  
Niveles didácticos. Comparación de medias entre los grupos experimental y control.

CICLO	SUBAREA APRENIDI.	GRUPO	N. <sup>o</sup> IND.	$\bar{X}$	DESV. STAN.	ERR. STAN.	T	G.L.	PROB.
INICIAL	«Logica»	Exp.	284	4,32	1,31	0,07	1,30	454	.193
		Cont.	172	4,16	1,21	0,09			
	«Numer.»	Exp.	284	20,95	4,65	0,27	5,53	454	.000
		Cont.	172	18,44	4,76	0,36			
	«Medid.»	Exp.	284	2,68	1,28	0,07	4,65	454	.000
		Cont.	172	2,11	1,25	0,09			
	«Geom.»	Exp.	284	4,61	1,25	0,07	3,05	454	.000
		Cont.	172	4,23	1,40	0,10			
MEDIO	«Logica»	Exp.	283	6,50	1,42	0,08	7,49	433	.000
		Cont.	152	5,38	1,58	0,12			
	«Numer.»	Exp.	283	28,80	4,95	0,29	7,29	433	.000
		Cont.	152	19,90	6,20	0,50			
	«Medid.»	Exp.	283	6,75	2,70	0,16	4,99	433	.000
		Cont.	152	5,42	2,55	0,20			
	«Geom.»	Exp.	283	4,22	1,95	0,11	6,65	433	.000
		Cont.	152	3,00	1,54	0,12			
SUPERIOR	«Numer.»	Exp.	204	10,39	4,71	0,33	2,50	372	.013
		Cont.	170	9,10	5,28	0,40			
	«Fun. Pol Prop.»	Exp.	204	3,34	2,23	0,15	2,98	372	.003
		Cont.	170	2,67	2,10	0,16			
	«Medid.»	Exp.	204	1,19	0,75	0,05	3,80	372	.000
		Cont.	170	0,90	0,68	0,05			
	«Geom.»	Exp.	204	7,91	2,96	0,20	7,34	372	.000
		Cont.	170	5,78	2,56	0,19			

TABLA XI

Subáreas de aprendizaje matemático. Comparación de medias entre los grupos experimental y control.

Las conclusiones que se obtuvieron, pueden resumirse en que la media de las puntuaciones obtenida en cada ciclo y nivel didáctico respecto a los dos grupos era más elevada, en todos los casos, en el grupo experimental siendo dicha diferencia siempre significativa al nivel .000 De ello se derivó que el Sistema TOAM en la muestra de alumnos en la que se trabajó, dió lugar en el grupo experimental a un más alto grado de asimilación de conceptos, propiedades y vocabulario matemático; que supieran utilizar los mecanismos de operatividad con una mayor calidad de producto; que el alumno resolviera situaciones problema (real o abstractas) con mayor efectividad.

\* *Estudio de las puntuaciones en las diversas áreas cognoscitivas*

En la materia de Matemáticas de E.G.B. se presentan varias áreas de estudio las cuales se enumeran en la TABLA XI. En esta fase de la investigación se detectó las diferencias que pudieran haber en cada ciclo entre los grupos control y la experimental por medio de una contratación de medias. De los resultados estadísticos (TABLA XI), se derivó:

- Que la diferencia entre grupos y subáreas directamente conexionadas al contenido del Sistema TOAM era significativa.
- Que en «Lógica» (C. Inicial) la diferencia entre grupos no era significativa.
- Que en el área de Geometría –no tratada con E.A.O.– también constaban diferencias significativas entre los dos grupos.

## RELACIÓN TEST TOAM-PRUEBA OBJETIVA

En esta última fase de estudio se pretendió indagar la relación y paralelismo existente entre el Test TOAM y las Pruebas Objetivas. Recordemos que aquel indica al profesor el nivel (curso/mes) en que se encuentra el alumno en lo que respecta a sus conocimientos en aritmética; las Pruebas dan un valor cuantitativo que equivale al número de ítems bien solucionados.

A fin de detectar esta relación se buscó la correlación existente entre la puntuación del alumno en cada tema del TOAM y la alcanzada en cada grupo de ítems de la Prueba que correspondía a dicho tema. No obstante cabe indicar que no fue posible en todos los casos hacer esta correlación; en estos temas se realizó un análisis comparativo de medias (C.I.: tema 6; C.M. tema 3 y 4). A través de los resultados estadísticos (TABLA XII) se vió que las Pruebas y el Test TOAM presentaba una elevada correlación significativa excepto en los temas 3 y 4 de C.M. Era pues necesario hacer en estos dos temas un estudio posterior.

## CONCLUSIONES

De este estudio y atendiendo a la doble finalidad del mismo se obtuvieron conclusiones agrupadas en tres ámbitos:

TEMAS	C.I.		C.M.		C.S.	
	Corr.	Sig.	Corr.	Sig.	Corr.	Sig.
1.-Sist. Numer.	.452	.000	.517	.000	.172	.000
2.-Ad. Sus. has. 20	.366	.000	—	—	—	—
3.-Adic. Vert.	.147	.005	Dif. Sign. .001	—	—	—
4.-Sust. Vert.	.325	.000	Dif. Sign. .000	—	—	—
5.-Mul. Div. has. 100	.504	.000	—	—	—	—
6.-Mult. Vert.	Dif. no sig.		.438	.000	—	—
7.-División	—	—	.554	.000	—	—
8.-Oper. Ment.	.395	.000	.556	.000	.518	.000
9.-Ecuac. Ley. Fund.	.389	.000	.506	.000	—	—
10.-Unid. Medida	.289	.035	.600	.000	.439	.000
11.-Frac. Simples	—	—	.515	.000	.563	.000
12.-Frac. Dec. Porc.	—	—	.412	.000	.515	.000
13.-Num. Enteros	—	—	—	—	.339	.000
14.-Probl. Verbal.	.499	.000	.688	.000	.616	.000
15.-Divis. Factor. Pot.	—	—	—	—	.406	.000

TABLA XII

Relación de los resultados del Test  
Toam con la Prueba Objetiva en  
los distintos temas.

\* *Pruebas Objetivas*

- Las tres pruebas objetivas creadas para este trabajo y correspondientes a los tres ciclos de E.G.B. tuvieron respectivamente coeficientes elevados de fiabilidad (.89, .91 y .89)
- La validez: en este sentido las tres pruebas fueron positivas.
- Respecto a cada uno de los ítems:
- cada ítem es exponente de un objetivo del nivel correspondiente
- La dificultad en cada uno de los diferentes reactivos fue muy distinta. Los índices oscilaron del 2% al 98'7%. Las pruebas en este sentido eran heterogéneas como corresponde a una prueba de matemáticas de este tipo.
- En las 165 correlaciones realizadas entre la respuesta de cada ítem y la puntuación global de su prueba se obtuvieron valores significativos excepto en tres ítems (C.I.: n.27; C.M.: n.36; y C.S.: n. 34). Será preciso revisarlos en una próxima fase de estudio. No obstante cabe señalar que no se creyó que afectaran significativamente a la finalidad de este trabajo. La aportación de cada ítem a la fiabilidad global de la prueba es semejante. Las tres pruebas se presentaron equilibradas y homogéneas en este aspecto.

\* *Estudio comparativo entre el grupo experimental y el control*

- No se manifestaron diferencias significativas respecto al factor «tiempo» de realización de las pruebas entre los dos grupos de estudio.

- El grupo experimental mostró mejores puntuaciones en la Prueba Objetiva que el grupo control.
- La dispersión de la «Puntuación global» en las Pruebas era más reducida en los grupos experimentales.
- Respecto a los diferentes niveles didácticos de cada Ciclo, la media de las puntuaciones fue más elevada en todos los casos en el grupo experimental que en el grupo control a un grado de significación de .000
- En el ámbito de áreas cognoscitivas de estudio en Matemáticas también se mostraron diferencias significativas, excepto en el área de Lógica del Primer Ciclo, entre ambos grupos.

\* *Relación Test TOAM-Prueba Objetiva.*

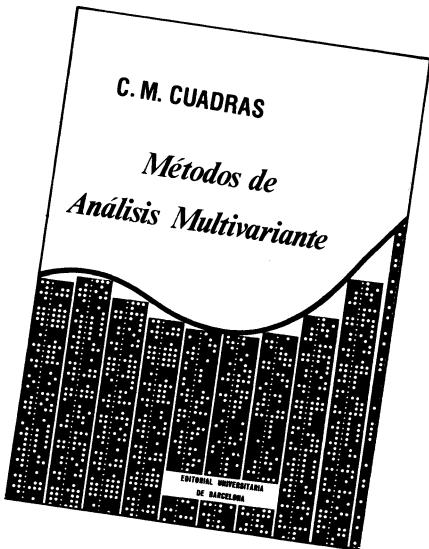
- Las correlaciones entre la puntuación media del alumno en el Test. TOAM y la Prueba Objetiva fue significativa (C.I.: .61; C.M.: .43; y C.S.: .72 siendo  $p$ . .000).
- Igualmente se presentaron elevadas y significativas las correlaciones entre las puntuaciones obtenidas por los alumnos en cada uno de los temas del Sistema TOAM y los puntos obtenidos en los items correspondientes en la Prueba. Fueron excepciones los temas 3 y 4 en el Ciclo Medio.

# Colección MATEMÁTICAS

## Métodos de análisis multivariante (C.M. Cuadras)

La obra recoge parte de las enseñanzas impartidas por el autor en diversas asignaturas, seminarios y cursos de doctorado en Facultades de Matemáticas y Biologías, así como su experiencia como analista de aplicaciones científicas en el Laboratorio de cálculo de la Universidad de Barcelona.

La obra va dirigida a estudiantes de segundo ciclo, postgraduados y profesionales en Informática, Biología, Medicina, Pedagogía, Psicología... etc., así como a estadísticos y matemáticos que deseen tener una visión apli-



cada de este tipo de análisis, sin renunciar a sus fundamentos teóricos.

Recoge al final una relación comentada de «packages» y programas de Análisis Multivariante.  
(P.V.P.: 2.400 pts.).

### Otras obras del autor y su equipo

C.M. Cuadras: *Problemas de Probabilidades y Estadística (Probabilidades)* (P.V.P.: 1.600 pts.).

C.M. Cuadras: *Problemas de Probabilidades y Estadística (Inferencia Estadística)* (P.V.P.: 1.600 pts.).

C.M. Cuadras y otros: *Fundamentos de probabilidad en Bioestadística* (P.V.P.: 1.600 pts.).

C.M. Cuadras y otros: *Ejercicios de Bioestadística* (P.V.P.: 550 pts.).

Los encontrará en los distribuidores de P.P.U.

**PROMOCIONES Y PUBLICACIONES UNIVERSITARIAS**

C/. Nicaragua, 100, 7.<sup>o</sup> 1.<sup>a</sup>, 08029-BARCELONA

Teléf. (93) 239 91 37

## Proyecto E.A.O. -TOAM Estudio piloto sobre variables asociadas al rendimiento

por  
Flor Cabrera  
Juana Gómez

### INTRODUCCIÓN

El presente estudio se enmarca dentro de los trabajos de investigación en torno al *Proyecto E.A.O.*, cuyas características se exponen en el trabajo del Dr. Benito Echeverría.

Se trata de un estudio meramente piloto, con el cual se pretende iniciar una línea de trabajo e investigación que nos acerque al conocimiento y descripción de variables psicopedagógicas relacionadas con las conductas de aprendizaje propuestas en las secuencias instrutivas del programa. Su finalidad última es proporcionar al profesorado una mayor comprensión del origen de las dificultades del programa y, en consecuencia, orientar más eficazmente su práctica docente.

En una primera fase, nuestro trabajo utiliza la metodología propia de los estudios correlacionales. Partiendo de las diferencias observadas en el rendimiento de los alumnos, intentamos constatar la relación entre variables tradicionalmente consideradas relevantes en el rendimiento aritmético y el aprovechamiento de los alumnos en los temas del sistema.

El centro de interés en un primer momento, aunque también se estudiarán variables aptitudinales, es la preocupación del profesorado y administradores del sistema en la posible influencia del nivel lector del alumno en lengua catalana, puesto que el programa está impreso en esta lengua y un elevado porcentaje de los alumnos que realizan la experiencia son de lengua materna y escolar castellana.

Los mismos resultados obtenidos en este estudio, (resultados no definitivos ni generalizables) nos indican la necesidad de continuar nuestro trabajo desde unas perspectivas metodológicas más cualitativas o más fenomenológicas, en el sentido en que es desarrollado por Stake, R.E. (1975) y Guba y E.G. y Lincoln, Y.S. (1982).

## MUESTRA

Dado el carácter piloto del presente estudio, la muestra de sujetos la constituye un aula escolar de 2.<sup>º</sup> de E.G.B. con 32 alumnos del centro experimental situado en el barrio de la Verneda de Barcelona —colegio «Els Horts»—. Se elige dicha aula por razones meramente de disponibilidad y facilidad en la recogida de la información que precisábamos. Su tipología ha sido ya descrita en el trabajo del Dr. Benito Echeverría.

## METODOLOGÍA

Se utiliza el método correlacional, tomando como:

- A) *Variables criterio*: el rendimiento de los alumnos en el total y en los distintos temas de aprendizajes del sistema T.O.A.M. propios de 2.<sup>º</sup> de E.G.B. (temas 1, 2, 3, 4, 5, 8, 9 y 14).
- B) *Variables explicativas*:

- \* *Comprensión lectora castellana: test standarizado* (Bartolomé y otros, 1985 a).
- \* *Comprensión lectora catalana: test standarizado* (Bartolomé y otros, 1985 b).
- \* *Los factores aptitudinales recogidos en las materias BAD YG-GRÁFICO B que incluye los siguientes dos factores de grupo y sus correspondientes subtests:*
  - *Factor de grupo: Madurez Intelectual Verbal, I.G.V.:*  
Subtest de Ordenes Verbales, O.V.  
Subtest de Aptitud para el Cálculo, N  
Subtest de Información, Inf.
  - *Factor de grupo: Madurez Intelectual No verbal, I.G.Nv.:*  
Subtest de Habilidad Mental No verbal, H.M. Nv.  
Subtest de Razonamiento con Figuras, R. L.  
Subtest de Atención-Observación, A.

## ANÁLISIS DE DATOS

Las variables estudiadas se han tratado primeramente a nivel correlacional, utilizando en su cálculo el paquete B.M.D.P. (Dixon W.J., 1983).

En el posterior estudio de regresión de las diferentes variables aptitudinales sobre la evaluación de los temas T.O.A.M. se ha utilizado el programa P 1 R (Múltiple Linear Regression), también del paquete B.M.D.P. En él analizare-

mos, por una parte, los coeficientes de regresión obtenidos ( $b_i$ ) y su nivel de significación (valores T) y, por otra, a partir de los respectivos coeficientes estandarizados, la contribución de cada variable independiente a la varianza explicada por la ecuación de regresión correspondiente (cálculo de  $B \times r$ ), Rodríguez, S. (1984).

El uso que hacemos aquí de la regresión, como es obvio por el mismo carácter de nuestros datos (limitado número de sujetos fundamentalmente), responde a una finalidad de técnica exploratoria de análisis de datos.

## RESULTADOS

La matriz de correlación expresada en la Tabla 1 permite constatar, tomando como criterio de riesgo de rechazo de la hipótesis nula un  $\alpha$  de 0,05, los siguientes resultados:

### 1. Correlaciones entre las variables criterio.

Las correlaciones entre los temas del sistemas presentan valores muy altos, oscilan entre 0,55 (t. 2 y t. 14) y 0,91 (t. 8 y t. 3). Las correlaciones de cada uno de los temas con el total son también muy elevadas y prácticamente muy parecidas, en torno al 0,90.

### 2. Correlaciones entre las variables criterio y las explicativas.

#### 2.1. Correlaciones entre los temas y las comprensiones lectoras.

La Comprensión Lectora Castellana mantiene correlaciones significativas con prácticamente todos los temas (excepto el tema 2 y 9), y con el total. En cambio, la Comprensión Lectora Catalana presenta coeficientes extremadamente muy bajos.

A destacar el incremento considerable de los coeficientes de correlación en el tema 1 y 14, sobre todo en lo que se refiere a la Comprensión Lectora Catalán, llegando incluso a un valor, en el tema 14, con un grado de significación de 0,09.

#### 2.2. Correlación entre los temas y los factores y subtests aptitudinales.

\* Con el total aptitudinal: todos los coeficientes se presentan significativos, destacando los valores elevados en los temas 1, 9 y 14.

\* Con los factores de grupo, I.G.V. y I.G.N.v.: tanto el total como cada uno de los temas, presentan correlaciones significativas con el factor I.G.V., alcanzado prácticamente en todos un grado de significación del 0,01. Mientras que con el factor I.G.N.v. los coeficientes son extremadamente bajos, sólo a destacar el valor de 0,33 para el tema 14, con un grado de significación del 0,06.

**TABLA 1.**  
**MATRIZ DE INTERCORRELACIONES**

	TEMAS SISTEMA T.O.A.M.										Comprensión lectora		VARIABLES APTITUDINALES								CALIFICACIONES		
	total	t.1	t.2	t.3	t.4	t.5	t.8	t.9	t.14	cast.	catl.	O.V.	N.	Inf.	N.v.	R.L.	At.	I.G.V.	N.v.	total	lenguas: cast.	catl.	Matemát.
total	—																						
t.1	.88	—																					
t.2	.72	.63	—																				
t.3	.94	.86	.64	—																			
t.4	.89	.75	.72	.84	—																		
t.5	.89	.71	.57	.82	.85	.81	—																
t.8	.92	.85	.79	.91	.85	.86	.82	—															
t.9	.95	.81	.76	.87	.86	.82	.89	.84	—														
t.14	.89	.76	.55	.81	.77	.80	.74	.84	.42	—													
Comp. l. cast.	.37	.40	.33	.38	.34	.36	.36	.28	.42	.58	—												
Comp. l. cat.	.16	.29	.12	.08	.11	.15	.05	.12	.30	.27	.38	—											
O.V.	.45	.59	.37	.46	.41	.34	.42	.38	.42	.44	.37	—											
N.	.54	.47	.54	.55	.53	.43	.55	.51	.50	.28	.14	.56	—										
Inf.	.37	.50	.32	.36	.14	.12	.27	.32	.38	.34	.38	.60	.36	—									
H.M.Nv.	.03	.07	-.04	.05	.05	.05	.01	-.03	.11	.37	.26	.30	.13	.11	—								
R.L.	.17	.09	.19	.15	.17	.09	.11	.15	.37	.29	.22	.40	.47	.20	.55	—							
At.	.18	.28	.11	.13	.03	.10	.10	.10	.31	.21	.32	.55	.49	.49	.52	.70	—						
I.G.V.	.59	.66	.52	.60	.47	.41	.54	.53	.56	.45	.40	.87	.76	.78	.25	.46	.61	—					
I.G.Nv	.23	.27	.18	.21	.13	.15	.16	.15	.33	.30	.27	.56	.50	.46	.77	.78	.88	.61	—				
total	.46	.53	.36	.47	.35	.32	.39	.59	.50	.44	.39	.81	.70	.67	.56	.69	.82	.90	.88	—			
l. cast.	.63	.61	.61	.56	.57	.57	.57	.63	.57	.43	.34	.59	.47	.47	.38	-.04	.07	.26	.60	.16	.43	—	
l. catl.	.67	.68	.65	.64	.56	.52	.67	.64	.55	.35	.32	.42	.51	.48	-.02	-.02	.20	.62	.17	.43	.71	—	
matem.	.92	.81	.74	.88	.86	.87	.90	.92	.77	.33	.09	.42	.41	.22	.09	.07	.06	.48	.15	.36	.66	.66	—

\* Con los subtests: Los valores de correlación más altos se encuentran con el subtest N, todos ellos significativos. Le sigue el subtest O.V. con correlaciones también significativas con todos los temas, destacándose el valor de 0,59 que presenta con el tema 1.

Las correlaciones con los subtests no verbales son prácticamente muy bajas, como eran de esperar a partir de los resultados del epígrafe anterior. No obstante, a destacar la correlación que con el tema 14 presenta el subtest R.L., con un valor significativo, y el subtest At. con un grado de significación del 0,07.

### 3. Correlaciones entre las variables explicativas:

#### 3.1. Correlaciones de las comprensiones lectoras.

\* La correlación entre las dos comprensiones lectoras es bastante elevada ( $r = 0,58$ ).

\* Los coeficientes que ambas comprensiones tienen con el total aptitudinal y los factores de grupo, repiten el patrón ya observado cuando nos referímos a los temas del sistema: valores más elevados y significativos con el total y el factor I.G.V., y coeficientes no significativos con el factor I.G.N.v.

\* En las correlaciones con los subtests, destacan las bajas correlaciones que ambas compresiones mantienen con el subtest N. Los coeficientes más significativos para las dos compresiones lectoras se encuentran con el subtest O.V. y el subtest Inf.

\* Con los subtests no verbales en general presentan correlaciones no significativas. Sólo a destacar el valor significativo de la Comprensión Lectora Castellana con el subtest H.M. Nv.

#### 3.2. Correlaciones de los subtests aptitudinales.

\* Como era de esperar, los valores de correlación entre los subtests que forman el factor de grupo I.G.V., por un lado, y entre los subtests que forman el factor I.G.N.v., por otro, son más elevados que los coeficientes entre subtests no pertenecientes a su factor de grupo.

\* El subtest At. destaca entre todos por ser el que presenta los valores de correlación más altos, todos significativos, con los otros subtests de la batería B.A.D.Y.G., y el subtest H.M.Nv. como el que menores valores presenta con los subtests no pertenecientes a su factor de grupo.

### 4. Correlación de las variables estudiadas con las notas escolares.

\* Con el total y cada uno de los temas: en general, se presentan elevados, oscilando entre los valores 0,52 y 0,92. Los valores más elevados se encuentran con la calificación en Matemáticas.

\* Con las compresiones lectoras: la Comprensión Lectora castellana mantiene correlaciones significativas con las dos lenguas, castellana y catalana, y con las matemáticas un coeficiente con un grado de significación del 0,06. Mientras que la comprensión lectora catalana no presenta correlación significativa alguna, destacando la prácticamente nula correlación que mantiene con matemáticas.

\* Con las aptitudes: correlación significativa con el total aptitudinal y con el factor I.G.V., pero no con el factor I.G.N.v., correlaciones elevadas con los subtests verbales (excepto el coeficiente no significativo entre la calificación en matemáticas y el subtest Inf.) y muy bajos con los subtests no verbales.

## 5. Análisis de la regresión múltiple.

### 5.1. Coeficiente $b$ y significación estadística.

Los datos de la Tabla 2 expresan para cada una de las variables criterio –temas del sistema– los coeficientes  $b$  y los respectivos valores correspondientes a las variables explicativas –las dos compresiones lectoras y los subtests aptitudinales–.

La observación de estos datos permite constatar: la única variable explicativa que presenta coeficientes  $b$  significativos se refiere al subtest numérico. Es en los temas 2, 3, 4, 8 y 9 donde este subtest alcanza significación. En el tema 5 se encuentra con un valor cercano al límite significativo. En las temas 1 y 14 su coeficiente baja considerablemente.

### 5.2. Coeficiente $\beta$ y contribución a $R^2$ .

En la Tabla 3 se recogen los coeficientes  $\beta$  de las variables explicativas y la contribución de las mismas ( $r_x \beta$ ) a la varianza explicada de la variable criterio por  $R^2$ . De los datos de la tabla se deduce:

\* Para los temas 2, 3, 4, 5, 8 y 9, el subtest Numérico es el que presenta mayor contribución a las  $R^2$  respectivas, en un porcentaje promedio del 63% (para el t. 2 un 63% de  $R^2$ ; para el t.3 un 58%; para el t. 4 un 63%; para el t.5 un 61%; y para el t.9 un 68%).

\* Para los temas 1 y 9 la varianza explicada de sus respectivas  $R^2$ , se encuentra más repartida, explicando aproximadamente el 60%.

\* Para el tema 1, es el subtest de órdenes verbales el que más contribuye, en un 44%. Le sigue en importancia el Numérico, con una contribución del 27%.

\* Para el tema 14 las variables que más contribuyen son el Numérico, 39%, Comprensión Lectora Castellana, 26% y Razonamiento Lógico, con un 23%.

TABLA 2. Coeficiente de determinación y coeficiente de regresión (con su correspondiente significación).

## TEMAS SISTEMA T.O.A.M.

	Variable criterio	tema 1 $r^2 = 0,482$		tema 2 $r^2 = 0,409$		tema 3 $r^2 = 0,476$		tema 4 $r^2 = 0,459$		tema 5 $r^2 = 0,309$		tema 8 $r^2 = 0,45$		tema 9 $r^2 = 0,37$		tema 14 $r^2 = 0,39$	
	Variables Explicativas	Coef. b	t	Coef.	t	Coef.	t	Coef.	t	Coef.	t	Coef.	t	Coef.	t	Coef.	t
	Compr. lect castellana	.351 -.012	0,939 -0,034	.252 -.089	0,831 -0,308	.572 -.457	1,174 -0,985	.299 -.031	0,709 -0,096	.556 -.086	1,306 -0,213	.655 -.463	1,698 -1,037	.210 -.056	0,477 -.056	.298 .117	0,993 .308
Madurez Intelectual Verbal (I.G.V.)	O.V. N. Inf.	.442 .284 .199	1,566 1,310 0,734	.064 .398 .213	0,280 2,255 0,923	.287 .695 .247	0,780 2,454 0,981	.439 .636 .126	1,373 2,586 -0,411	.240 .429 -.243	0,720 1,732 -0,788	.292 .675 .104	0,825 2,476 0,306	.204 .636 .330	0,538 2,171 0,901	.028 .320 .211	0,095 .377 .762
Madurez Intelectual no Verbal (I.G.N.v.)	(H.M.N.v) R.L. At.	-.049 -.505 .089	-0,161 -1,188 0,209	-.122 .204 -.390	-0,489 0,591 -1,135	.093 -.140 .565	0,232 -0,253 -1,020	.077 .110 .761	0,222 0,228 -1,582	-.058 -.371 .018	-0,167 -.763 0,039	-.024 -.228 -.389	-0,063 -0,428 -0,731	-.077 .160 -.667	-0,185 0,280 -1,164	-.200 .442 -.172	-0,606 .971 -.380

TABLA 3: Coeficientes de regresión estandarizados y proporción de varianza explicada por cada variable explicativa

	Tema 1		tema 2		tema 3		tema 4		tema 5		tema 8		tema 9		tema 14	
	B	r.B														
C. Lectura Catellana	0,203	0,008	0,192	0,06	0,255	0,10	0,157	0,05	0,326	0,12	0,310	0,11	0,099	0,03	0,233	0,10
C. Lectura Catalana	-0,007	0,02	-0,064	-0,01	-0,194	-0,02	-0,019	-0,01	-0,048	-0,01	-0,209	-0,01	-0,026	0,01	0,065	0,02
Q. V.	0,357	0,21	0,068	0,03	0,179	0,08	8,320	0,14	0,196	0,07	0,193	0,08	0,135	0,05	0,024	0,01
N.	0,267	0,13	0,491	0,26	0,504	0,28	0,539	0,29	0,408	0,19	0,519	0,29	0,487	0,25	0,305	0,15
Inf.	0,156	0,08	0,210	0,08	0,210	0,08	-0,090	-0,01	-0,194	-0,02	0,067	0,02	0,211	0,07	0,176	0,06
H.M.Nv.	-0,034	-0,02	-0,109	-0,01	0,049	0,01	0,047	0,01	-0,040	-0,01	-0,014	-0,01	-0,042	0,01	-0,137	-0,02
R.I.	-0,287	-0,03	0,153	0,03	-0,062	-0,01	0,057	0,01	-0,214	-0,02	-0,106	-0,01	0,074	0,01	0,254	0,09
At.	0,059	0,01	-0,340	-0,04	-0,289	-0,04	-0,455	-0,01	0,013	0,01	-0,211	-0,02	-0,360	-0,04	-0,116	-0,03
	$r^2 = .48$		$r^2 = .41$		$r^2 = .48$		$r^2 = .46$		$r^2 = .31$		$r^2 = .45$		$r^2 = .37$		$r^2 = .39$	

### 5.3. Análisis complementarios.

Se realizaron además otros dos análisis de regresión múltiple para complementar la exploración de los datos:

\* Considerando como variables explicativas únicamente los 6 subtests del B.A.D.Y.G., el patrón de resultados es prácticamente igual que el comentado anteriormente.

En efecto, el subtest Numérico es el que aparece de modo constante con significación estadística, excepto en los temas 1, 5 y 14. Es a destacar que la variable Numérica aumenta su contribución a la  $R^2$  en todos los temas, salvo para el tema 1 que permanece prácticamente igual; en este último tema aumenta, en cambio, la contribución de la variable Ordenes Verbales. Para el tema 14, aumenta también la contribución del Razonamiento Lógico y aparece cierta relevancia del subtest de Información.

En la tabla 4 se expresan las contribuciones más significativas a  $R^2$  de los subtests aptitudinales, cuando se realiza la regresión con o sin las dos pruebas de comprensión lectora.

\* Tomando como variables explicativas sólo los 3 subtests verbales del B.A.D.Y.G., los coeficientes significativos se presentan sólo para el subtest Numérico, excepto en los temas 1 y 5. Respecto a la contribución de las variables a sus respectivas  $R^2$ , destaca para todos los temas –exceptuando de nuevo el tema 1– el subtest Numérico, muy por encima de los demás subtests. En el tema 1, la mayor contribución corresponde al test de Ordenes Verbales, con un 51% de  $R^2$ ; le sigue el test de Informaciones con un 27%, y por último el Numérico con un 22%.

En la tabla 5 se recogen los respectivos coeficientes de regresión estandarizados y su correspondiente contribución a la correlación múltiple al cuadrado.

## DISCUSIÓN

Los resultados obtenidos en la matriz de correlaciones evidencia algunos aspectos de interés para el objetivo de nuestro estudio:

a) La diferencia ostensible entre los valores de correlación que las dos comprensiones lectoras presentan con el rendimiento en los temas del sistema: la Comprensión Lectora Castellana presenta valores de correlación estadísticamente significativos con la mayoría de los temas, mientras que la comprensión Lectora Catalana presenta en general valores muy bajos.

Añadimos a estos resultados las correlaciones significativas que presentan la Comprensión Lectora Castellana con las calificaciones escolares, y no así la Comprensión Lectora Catalana.

Podríamos aventurar como una posible explicación de estos resultados, el

TABLA 4. Contribución de las variables aptitudinales de  $R^2$ , en función de si se introducen o no las dos Comprensiones Lectoras.

	Tema 1		tema 2		tema 3		tema 4		tema 5		tema 8		tema 9		tema 14	
	x	xx	x	xx												
$R^2$	.48	.45	.41	.39	.47	.44	.47	.45	.31	.26	.45	.40	.37	.37	.39	.34
O.V.	.44	.53			.16	.20	.30	.31	.23	.31	.17	.20	.68	.70	.39	.50
N.	.27	.29	.63	.66	.58	.62	.63	.65	.61	.68	.64	.78	.19	.20		.22
Inf.	.17	.20	.15	.21	.16	.17									.23	.29
H.M.Nv.																
R.L.																
At.																

x Como variables explicativas se han introducido las variables aptitudinales y las dos pruebas de comprensión lectora.

xx Se han introducido únicamente las variables aptitudinales (en ambos casos, indicamos sólo los valores que presentan una contribución mayor al 15% de  $R^2$ ).

TABLA 5. Coeficientes de regresión estandarizados y su correspondiente proporción de varianza, cuando sólo se introducen en la regresión las aptitudes verbales

	TEMAS SISTEMA T.O.A.M.															
	Tema 1		tema 2		tema 3		tema 4		tema 5		tema 8		tema 9		tema 14	
	B	r.B	B	r.B	B	r.B	B	r.B	B	r.B	B	r.B	B	r.B	B	r.B
O.V.	.357	.21	.021	.01	.141	.06	.281	.13	.229	.08	.155	.06	.069	.03	.098	.04
N.	.189	.09	.480	.26	.436	.24	.450	.24	.355	.15	.461	.21	.431	.23	.389	.20
Inf.	.219	.11	.136	.05	.122	.04	-.196	-.03	-.142	-.02	.012	.01	.128	.04	.190	.07
$R^2$	.41		.31		.35		.32		.21		.32		.29		.30	

papel de lengua instrumental que la lengua castellana ocupa en el aprendizaje de los sujetos de la muestra. Es un hecho constatado en muchas otras investigaciones la relación entre el nivel lector y el rendimiento escolar, fundamentalmente en los bajos niveles escolares. (Lapp. D. y Flood Y. 1978).

b) Llama la atención el incremento notable de la correlación que presenta la Comprensión Lectora Catalana con los temas 1 y 14, (con un grado de significación en torno al 0,07, muy superior al resto de los temas). En la Comprensión Lectora Castellana, nos encontramos también con valores mayores en estos dos temas. Añadimos a estos resultados, la elevada correlación que tienen entre sí las dos comprensiones lectoras, 0,58, que es el más elevado entre los coeficientes calculados para estas dos variables.

Los resultados que comentamos ¿no ponen en evidencia un proceso mental común entre ambas comprensiones lectoras? ¿Proceso además exigido en la resolución de los ejercicios de los temas 1 y 14 pero no en los otros?. Un somero análisis del contenido de los temas nos indica el mayor contenido de órdenes verbales que éstos implican en sus ejercicios. El tema 1 presenta órdenes verbales cuya comprensión es imprescindible para la resolución del ejercicio; el tema 14 presenta problemas verbales de razonamiento aritmético. En los demás temas el contenido verbal es prácticamente nulo.

Una aportación que ayudaría a aclarar estos interrogantes sería la réplica de esta experiencia en alumnos con lengua catalana dominante en el aprendizaje; ¿se invertiría o, por lo menos, cambiaría el perfil registrado con respecto a ambas comprensiones lectoras?

Otra posible aportación significativa podría ser realizar la experiencia en niveles escolares superiores donde los temas 1 y 14 presentan mayor contenido verbal en sus ejercicios, además de una previsible mayor variabilidad del nivel de los sujetos en ambas comprensiones lectoras.

c) Con respecto a los variables aptitudinales estudiadas, los resultados encontrados ponen de manifiesto, en una primera apreciación, que las elevadas correlaciones que mantienen el factor de grupo I.G.V. con todas las variables que implican rendimiento escolar son coincidentes con el autor de la batería B.A.D.Y.G. al identificar este factor con el «verbal educativo» de Vernon o la «inteligencia cristalizada» de Cattell. Para nuestro objetivo, los resultados más significativos aparecen cuando estudiamos aisladamente los subtests.

d) Se constata la alta incidencia del subtest Numérico en todos los temas del sistema. Teniendo presente que el tipo de tarea requerido en este subtest es cálculo mental, ello confirma la presencia de dicha habilidad en los ejercicios de aprendizaje propuestos en los temas del sistema.

e) El subtest de Órdenes Verbales también presenta correlación significativa con todos los temas, aunque con un incremento notable en el valor de correla-

ción con el tema 1. Este hecho es coincidente con el contenido verbal que anteriormente comentábamos en los ejercicios de este tema, que trata precisamente de órdenes verbales.

f) Los resultados analizados quedan aún más clarificados con los complementarios análisis de regresión múltiple efectuados. El subtest Numérico aparece como la aptitud más significativa y explicativa en los temas 2, 3, 4, 5, 8 y 9. Para el tema 1, el subtest de Órdenes Verbales aparece en primer lugar, seguido del Numérico; en cuanto al tema 14, son las variables Numérica, Comprensión Lectora y Razonamiento Lógico, respectivamente, las de mayor valor explicativo.

Cabe también destacar que en el tema 5 ninguna de las variables explicativas usadas en el estudio obtiene significación en los sucesivos análisis de regresión efectuados. Por otra parte, el  $R^2$  obtenido con respecto a este tema es considerablemente inferior a los hallados en otros temas del programa, alcanzando únicamente un valor de 0,31 cuando se introducen todas las variables en el análisis de regresión. Estos resultados ponen en cuestión la representatividad de los subtests utilizados en el estudio, en relación a las aptitudes mentales que subyacen en la base de los ejercicios del tema 5.

#### REFERENCIAS BIBLIOGRÁFICAS

- BARTOLOMÉ, M. (1985 a), «Test de Lectura», en M. Bartolomé y otros, *Batería de pruebas de lenguaje: Final del ciclo inicial*, CEAC, Barcelona.
- BARTOLOMÉ, M. (1985 b), «Test de Lectura», en M. Bartolome i altres, *Bateria de proves de llenguatge: Final de cicle inicial*, CEAC, Barcelona.
- DIXON, W.J. (Ed.) (1983), *B.M.D.P.: Biomedical Computer Programs*. University of California Press, Berkeley.
- ECHEVERRIA, B. (1985) «Proyecto E.A.O.-TOAM. Evaluación del rendimiento aritmético», *Revista de Investigación Educativa*, Vol. 3, 6.
- GUBA, E.G., Lincoln, Y.S. (1892), *Effective Evaluation*. Jossey Bass, Publ., California.
- JUSTE H.C. (1979), *BADYG*, CEPE, Madrid.
- LAPP, D. Flood, Y. (1978), *Teaching reading to every child*. Mc. Millan Publ. New York.
- RODRIGUEZ, S. (1984), «Precisiones y limitaciones explicativas en los métodos correlacionales. Alternativas metodológicas», *Revista de Investigación Educativa*, 4, 103-132.
- STAKE, R.E. (Ed.) (1975), *Evaluating the arts in education: A responsive approach*. Merrill, Columbus, Ohio.

# **DISTRIBUIDORES DE P.P.U. PROMOCIONES Y PUBLICACIONES UNIVERSITARIAS**

**C/. Nicaragua, 100, 7.<sup>o</sup> 1.<sup>a</sup>, 08029-BARCELONA  
Teléf. (93) 239 91 37**

## **DELEGACION ANDALUCIA**

DERRI-Distribuidora Sevillana  
Pol. Ind. Carretera Amarilla Serra y Pickman,  
parcela 159, nave 2H - Tel. (954) 51 67 77  
41007 - SEVILLA  
(Sevilla-Cádiz-Huelva-Badajoz)

Sr D. Salvador ROMERO  
Alvarez, 6 - Tel. (952) 22 49 68  
29008 - MALLAGA  
(Málaga-Almería-Córdoba-Jaén)

## **DELEGACION ARAGON**

ARDIS-Aragonesa de Distribución  
Hermanos Pinzón, 4 - Tel. (976) 34 43 43 - 27 54 96  
50010 - ZARAGOZA  
(Zaragoza-Huesca-Navarra-Soria-Logroño-Teruel)

## **DELEGACION ASTURIAS**

PRINCIPADO-LIBROS  
Monte Gamonal, 45 bajos - Tel. (985) 29 44 24  
33012 - OVIEDO  
(Oviedo-Santander)

## **DELEGACION BALEARES**

PALMA-DISTRIBUCIONES  
Colubi, 4 - Tel. (971) 28 94 21  
PALMA DE MALLORCA  
(Mallorca-Menorca-Ibiza)

## **DELEGACION CASTILLA**

DISTRIBUIDORA GENERAL DE LIBROS TAPIA  
Campomanes, 5 - Tel. (91) 470 01 00 - 470 00 12  
28013 - MADRID  
(Madrid-Toledo-Ciudad Real-Cuenca-Guadalajara-Salamanca-Palencia-Burgos-Segovia-Zamora-León-Valladolid)

## **DELEGACION CATALUÑA**

SERVICIO UNIVERSIDAD  
Medas, 8-10 - Tel. (93) 219 44 58  
08023 - BARCELONA  
(Barcelona-Tarragona-Lérida-Gerona)

## **DELEGACION GALICIA**

ESTEVEZ-DISTRIBUCIONES  
San Salvador de Poio-Veiro - Tel. (986) 85 71 01  
Apartado, 412  
PONTEVEDRA  
(Pontevedra-Lugo-Orense-La Coruña)

## **DELEGACION CANARIAS**

LEMUS-DISTRIBUCIONES  
Catedral, 29 - Tel. (922) 25 14 61  
LA LAGUNA (Tenerife)

## **DELEGACION MADRID**

REYDIS - LIBROS  
Huesca, 29 - Tel. (91) 279 18 24  
28020 - MADRID

## **DELEGACION LEVANTE**

DISVESA  
San Andrés, 4 - Tels. (96) 351 09 94 - 321 28 31  
46002 - VALENCIA  
(Valencia-Castellón)

## **JOSE GUILLEM ROBLES**

Espronceda, 35 - Tel. (965) 20 21 07  
ALICANTE  
(Alicante-Murcia-Albacete)

## **DELEGACION PAIS VASCO**

BLUME - DISTRIBUCIONES  
Iturrebide, 106 - Tel. (94) 433 08 77  
48006 - BILBAO  
(Vizcaya-Guipúzcoa-Alava-Santander)

## **DELEGACION PORTUGAL**

LITEXA - PORTUGAL  
Av. Rainha Danelia, 22 - Tel. 79 47 93  
LISBOA-5

## **DELEGACION AMERICA DEL NORTE**

LIBROS Y EDITORIALES, S.A.  
Av. Progreso, 202 - Tel. 5 16 42 61  
MEXICO-18 D.F.

## **DELEGACION AMERICA DEL SUR**

EDITORIAL LEGASA  
Rawson, 17 "A" - Tel. 983 24 92 - 983 24 94  
BUENOS AIRES (Argentina)

## **EDITORIAL REVERTE VENEZOLANA, S.A.**

Apartado de Correos, 68685  
Tels. 572-44 68 - 572-66 70  
CARACAS 1062-A (Venezuela)

## **EDITORIA REVERTE LTDA.**

Avenida do Exercito, 49 - C.P. 23001  
Tel. 284-52 44  
20910 RIO DE JANEIRO (Brasil)

## **EDITORIAL BLUME ECUADOR**

Juan León Mera, 775  
P.O. Box, 280-A - Tel. 54 03 80  
QUITO (Ecuador)

## Proyecto E.A.O – TOAM Evaluación del rendimiento aritmético

por  
Benito Echeverría

### CONTEXTO GENERAL

El Departament d'Ensenyament de la Generalitat de Catalunya ha puesto en funcionamiento diversos proyectos de informática aplicada a la educación, especificados en el gráfico 1 con el número de centros que participan.

El *Proyecto EAO* se ha iniciado en el curso 84-85 y al igual que el *Proyecto LOGO* va dirigido a alumnos de EGB

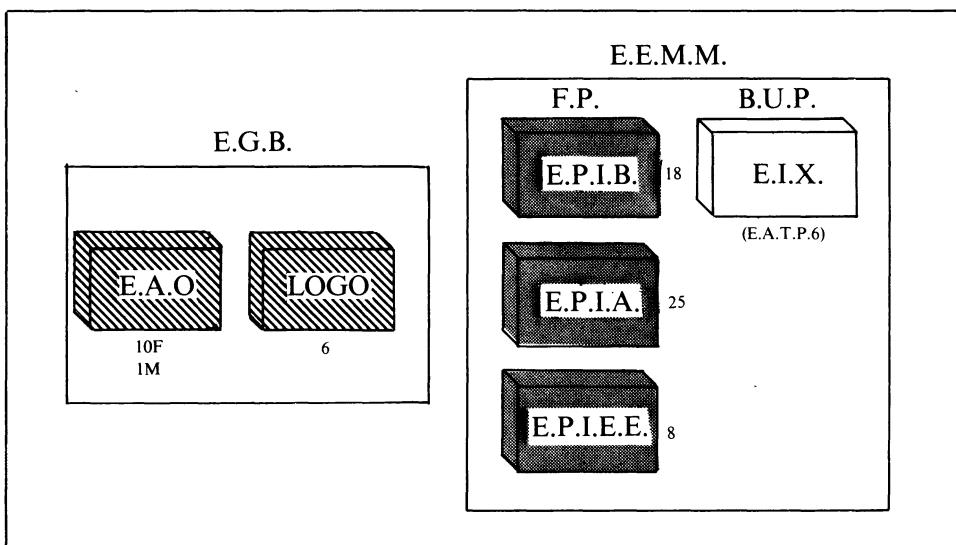


Gráfico 1: Proyectos y experiencias de introducción de la informática en la enseñanza (Generalitat de Catalunya).

Todos están coordinados por la «Comissió d'Informàtica del Departament d'Ensenyament», (Ordre 485; D.O.G 14-XI-84), que trabaja en el diseño de un programa general de actuación y sus tareas prioritarias son las de «coordinación, seguimiento y evaluación de las experiencias que se vienen realizando con el fin de conseguir un conocimiento real tanto en el aspecto estrictamente técnico pedagógico, como en la formación del profesorado en la ciencia informática, ya sea a nivel de contenidos, como de metodología».

De esta Comisión dimana el encargo de un diseño general de investigación y la evaluación de resultados del primer año de experiencia del *Proyecto EAO*, que se nos hizo en Noviembre de 1984.

Sobre este último aspecto centraremos nuestra exposición, tras un breve resumen del origen y características del sistema, sobre el que se sustenta el proyecto.

### Orígenes del Proyecto

Los orígenes de este proyecto se remontan a la década de los setenta, cuando el Centro de Tecnología Educativa (CET) de Tel Aviv realiza un exhaustivo análisis sobre al efectividad de los diferentes sistemas CAI, dirigido por el Dr. Y. Peless. De él se deriva que el sistema desarrollado por el equipo de la Universidad de Stanford (Suppes, P, et alt. 1972 y Macken, E, et alt. 1976) ofrece las máximas garantías para colaborar en la solución de alguno de los grandes problemas que presentan los alumnos asiáticos y africanos, inmigrados a Israel.

Uno de ellos, probablemente el más grave, se producía en el campo de la Aritmética, donde sólo el 25% de los alumnos procedentes de familias socioeconómicamente desfavorecidas lograban pasar de grado, frente al 72,2% de los nacidos en núcleos familiares con un alto nivel socioeconómico, que superaban el currículum establecido (Minkovich, A, et alt. 1977).

Los resultados de este estudio y el carácter básico de la materia condujeron al CET a dirigir sus primeros esfuerzos a la adaptación del programa de Aritmética de Stanford (CCC: Computer Curriculum Corporation). Los mayores cambios se introdujeron en el tema de «problemas verbales», donde la profesora de la Universidad de Haifa –Nesher, P.– ha aportado una valiosa colaboración con estudios sobre la naturaleza de estos problemas (1980), sobre los aspectos verbales (1976) y semánticos (1977 y 1982), relacionados con los mismos y sobre los estilos cognitivos que éstos requieren (1978).

Una vez diseñado y experimentado el «courseware» de Aritmética, elegido porque «los resultados obtenidos con los métodos anteriores no eran satisfactorios, se trataba de un temario obligatorio para todos los escolares y estaba bien definido en cuanto a contenidos y requisitos» (Osin, L, 1981, pág 469), el paso siguiente fue elaborar el «hardware» que sirviese a los fines curriculares y pedagógicos perseguidos. (Véase sus características en *Revista de Investigación Educativa*. Vol 2, N.º 3, págs. 62-64).

Entre estos últimos, el equipo del CET, compuesto en la actualidad por más de cien técnicos con dedicación exclusiva al centro, mantiene una postura claramente a favor de la enseñanza individualizada que resume su director con estos términos: «No todos los niños son capaces de progresar al ritmo requerido» (Osin, L, 1981, pág. 469), pero todos tienen «derecho a que se les provea de la oportunidad y de los medios para desarrollar completamente su potencial» (Osin, L, 1984, pág. 418).

De esta filosofía, mantenida por el CET desde que hizo la primera experiencia en Netivot -1977- con el «courseware» de Aritmética, participan las sucesivas ampliaciones en: a) *Drill and practice* (inglés como segundo idioma); b) *Tutorial* (electricidad); c) *Simulación* (geometría); d) *Problem solving* (LOGO) y e) *Lenguajes informáticos*.

De todos ellos los más experimentados son los de «ejercicio y práctica» y el más elaborado el «courseware» de Aritmética. Con él se ha iniciado el *Proyecto EAO*, tras previo estudio sobre las concordancias y discrepancias con las «Orientaciones i Programes» del Departament d'Ensenyament.

### Sistema TOAM

La estructura básica de este sistema, que así se llama respondiendo a las iniciales hebreas de «Test y Ejercitación Ayudado por Computador», engloba en su seno el Modelo de Conocimiento (MC), el Modelo del Estudiante (ME) y el Modelo Pedagógico (MP).

Para facilitar una mejor comprensión del proceso de investigación, que comentaremos después, resumimos sus características y remitimos al artículo de Osin, L, (1984) para una mayor profundización. De él se ha tomado la terminología.

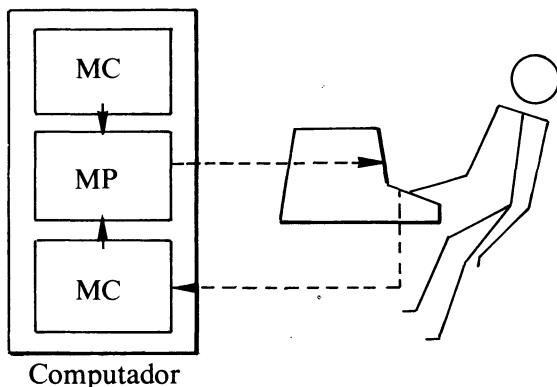


Gráfico 2: Estructura básica del sistema TOAM

En síntesis, podemos describirlo como un proceso en «feed-back» donde el modelo pedagógico (MP) decide qué tipo de información de MC ha de remitirse al estudiante en función de su perfil aptitudinal y ritmo de trabajo. Este procesa la información, contesta y su respuesta pone al día el modelo del estudiante (ME), que a su vez actualiza el modelo pedagógico (MP). A tenor de ella, el MP toma una nueva decisión y se reinicia el diálogo.

### a) Modelo de conocimiento (MC)

El modelo básico de casi todos los «courseware» de TOAM, y concretamente el de Aritmética, viene delimitado por tres dimensiones: a) Los *temas*; b) Los *niveles* de cada uno de ellos, y c) Los *ejercicios y secuencias* para cada tema y nivel.

Así, el «courseware» de Aritmética se compone de quince temas con niveles que fluctúan entre 10 y 79 (Véase Gráfico 3). Estos números indican los niveles ideales de curso y mes de trabajo, correspondientes a las orientaciones curriculares. Las decenas sirven para indentificar el curso escolar y las unidades el mes de trabajo, admitiendo años escolares de diez meses de duración. Por ejemplo, el nivel 58 contiene ejercicios que idealmente deberían dominar los alumnos de 5.<sup>º</sup> de EGB en el 8.<sup>º</sup> mes de clase.

Niveles:	10	20	30	40	50	60	70	80	RANGOS
TEMAS									
1									10-76
2									10-29
3									25-47
4									26-49
5									20-37
6									33-54
7									37-59
8									25-79
9									17-79
10									40-72
11									40-79
12									50-79
13									50-79
14									20-69
15									44-79

GRÁFICO 3: Courseware de Cálculo Aritmético, dividido en temas y niveles.

**b) Modelo del estudiante (ME)**

A través de él se puede conocer el bagaje de conocimientos del alumnado y sus respectivos ritmos de trabajo, ya que después de cada sesión los docentes no sólo reciben un informe detallado de aspectos administrativos (nombres del alumnado, clases,...). sino también del:

- a) Número de ejercicios presentados a cada alumno.
- b) Número de los mismos contestados correctamente.
- c) Porcentaje de errores debidos a falta de tiempo.
- d) Número de sesiones de trabajo realizadas.
- e) Nivel alcanzado en cada uno de los temas en los que les corresponde trabajar de acuerdo a su nivel medio.
- f) Temas, niveles y subniveles en los que un alumno no progresó o retrocede con fácil posibilidad de identificación, bien a través del computador o mediante una guía de patrones de ejercicios que disponen los docentes.

Toda esta información es convenientemente procesada en el modelo pedagógico y, como es patente, facilita la intervención individualizada de los docentes sobre la base de este diagnóstico previo.

**c) Modelo pedagógico (MP)**

Por razones de brevedad pasamos por alto el detenido y pormenorizado tratamiento que el computador realiza de la información procedente del resto de modelos y sólo esbozamos las fases de test y de ejercitación que definen el modelo.

**c1) Test**

Este se realiza a lo largo de las doce primeras sesiones de 10 minutos que el alumno trabaja con el computador y se acostumbra a repetir cada inicio de curso, para identificar los niveles de aprendizaje del alumnado en cada uno de los temas.

Funciona como los denominados «tailored testing» (Lord, F.L., 1970 y 1980), de tal forma que en la primera sesión se fija para el alumno un nivel de aprendizaje por debajo del programa que cursa y, a partir de ahí, siempre que éste trabaje por debajo del tanto por cien de éxitos esperados, baja el nivel de los ejercicios que se presentan a continuación. Al revés ocurre, cuando trabaja por encima.

El tanto por ciento de éxitos requeridos entre la sesión primera y cuarta es del 50% y entre la sesión quinta y duodécima del 67%, pero mientras entre la sesión quinta y octava el progreso y retroceso es global, de la sesión novena a la duodécima se realiza independientemente para cada tema del «courseware».

## c2) Ejercitación

Acabado el test, comienza esta fase que en Catalunya se realiza en dos períodos semanales de 20 minutos a lo largo de todo el curso. Es decir, cada vez que el alumno asiste a la sala de ordenadores realiza dos sesiones de 10 minutos, habitualmente conocidas con el nombre de «lecciones». En un curso normal el tiempo de ejercitación de estos alumnos vendrá a ser aproximadamente de 20 horas/año.

En esta fase el progreso se valora en términos de dominio y por eso el computador presenta al alumno entre 5 a 8 ejercicios por cada una de las escaleras de niveles-tema, según los criterios que se especifican en Cuadro 1. Siempre que se considere oportuno, se pueden cambiar los requerimientos de dominio que a continuación especificamos.

GRITERIOS DE EXITO				
Grupo de ejercicios	Grupo de 5	%	Grupo de 8	%
1	5/5	100	8/8	100
2	4/5	80	7/8	87,5
3...	3/5	60	6/8	75

Cuadro 1: Criterios de dominio en la fase de ejercitación.

Es importante resaltar que los alumnos difícilmente repiten ejercicios seguidos de un mismo tema, dado el alto número de «patrones» de secuencias por «courseware» (1.400 en Aritmética), que traducidos en términos de ejercicios supone un conglomerado de aproximadamente 10.000 ejercicios diferentes. Su elección es al azar con ponderación correctiva, que tiene en cuenta entre otros aspectos las dificultades presentadas por el alumno y la importancia relativa del tema.

## PROCESO DE INVESTIGACIÓN

### Objetivos investigación. Curso 84-85.

En la evaluación de este primer curso de experiencia del *Proyecto EAO* se ha pretendido:

- 1) Realizar un extensa prospección de los dominios de cálculo aritmético poseidos por los alumnos de 2.<sup>º</sup> a 8.<sup>º</sup> de EGB, que participan temporalmente (Aula móvil) o de forma continuada en el Proyecto, *antes de iniciar la fase de ejercitación con el sistema*.

- 2) Inferir a partir de esta prospección los *promedios de progreso en cálculo aritmético*, que lograban habitualmente los alumnos de la población, a la que pertenece la muestra analizada.
- 3) Contrastar en Catalunya la hipótesis que preconizan los creadores del sistema, al prever un *progreso medio de un nivel por mes de ejercitación* con el «courseware» de cálculo aritmético.
- 4) Controlar este progreso mediante un diseño cuasiexperimental en determinados centros de un mismo barrio, en el que unos alumnos actúan como grupos experimentales (Ejercitación *con* el sistema) y otros, de similares características, lo hacen como grupo control (Ejercitación *sin* el sistema).

Para aminorar en la medida de lo posible las fuentes de invalidez del diseño, sobre todo las incontroladas y posibles, (Interacción de variables, interacción de selección, tratamiento experimental, reactividad,...), se ha ejercido un mayor control sobre este último objetivo y se ha creado para ello instrumentos externos de medida.

En el trabajo de la Dra. Inmaculada Bordas se informa más detenidamente sobre las características de las tres pruebas objetivas creadas «ad hoc», para evaluar los conocimientos adquiridos en matemáticas, al final de cada ciclo de la E.G.B. A través de ellas se ha pretendido:

- 5) Contrastar las diferencias existentes entre grupos control y experimentales en:
  - Proceso global de aprendizajes matemáticos.
  - Niveles didácticos: «Conceptos», «Habilidades», «Mecánica» y «Razonamiento».
  - Dominios de aprendizajes en las distintas áreas de contenido de la materia de matemáticas.

## Muestra

El primer y segundo objetivo de investigación se ha analizado a partir de las observaciones recogidas en el *aula móvil*, itinerante por las comarcas catalanas, y en las nueve *aulas fijas*, asentadas en otros tantos centros de EGB.

### a) Aula Móvil

En el curso 84-85 esta aula ha recorrido la comarca de Osona y en ella los alumnos sólo han realizado la fase de test.

Se trata de una típica comarca agropecuaria, con bastante densidad de población y donde la tierra se explota en régimen familiar, siendo muy escasos los asalariados y jornaleros.

Los 6.646 alumnos de 2.<sup>º</sup> a 8.<sup>º</sup> de EGB, que pasaron por esta aula, fueron elegidos a través de un muestreo opinático, según criterios fijados por el Departament d'Ensenyament, tras previo informe de la Inspecció d'Educació Bàsica.

Los principales criterios para elección fueron:

- \* Territorios *sin aulas fijas* en su demarcación.
- \* Mayoría de *alumnos catalano-parlantes*.
- \* Posibilidad de realizar la prueba en *centros públicos y privados*.

#### b) Aulas fijas experimentales

Como se puede derivar de la exposición de objetivos, la investigación para el curso 84-85 se ha centrado fundamentalmente en las observaciones recogidas en estas aulas, a partir de las cuales se describe la situación inicial del Proyecto y se contrastan las hipótesis subyacentes a los objetivos tercero, cuarto y quinto.

Estas aulas, ubicadas en todas las provincias de Catalunya, se asientan en los centros que participan de las dos fases (test y ejercitación) del *Proyecto EAO* y en algún caso su utilización es compartida con otros colegios cercanos.

La elección de los 6.125 alumnos de estas aulas también la realizó el Departament d'Ensenyament con el mismo sistema de muestreo que en el caso anterior, pero dentro de una población, cuya tipología de centros venía definida por las siguientes características:

- \* Escuelas *públicas*, situadas normalmente en el extrarradio de ciudades o poblaciones.
- Los de la provincia de Barcelona están en un barrio obrero de la capital («La Verneda») y en Terrassa, Hospitalet, Badalona, St. Boi. Los de Tarragona en Reus y en un barrio de la capital de la provincia. El de Lleida es un centro de Acción Especial y el de Girona un colegio recién estrenado.
- \* Colegios con mayoría de *alumnos castellano-parlantes*, del orden de un 85% en el Ciclo Inicial, un 81% en el Medio y un 74% en el Superior.
- \* Zonas de *bajo rendimiento académico*, especialmente en Matemáticas. La tendencia de «insuficientes» (Junio) en el último trienio se representa en el Gráfico 4, que muestra los niveles de un mismo año académico (lectura horizontal), el mismo nivel en diferentes años académicos (lectura vertical) y el mismo grupo de alumnos a lo largo de su escolaridad (lectura diagonal).

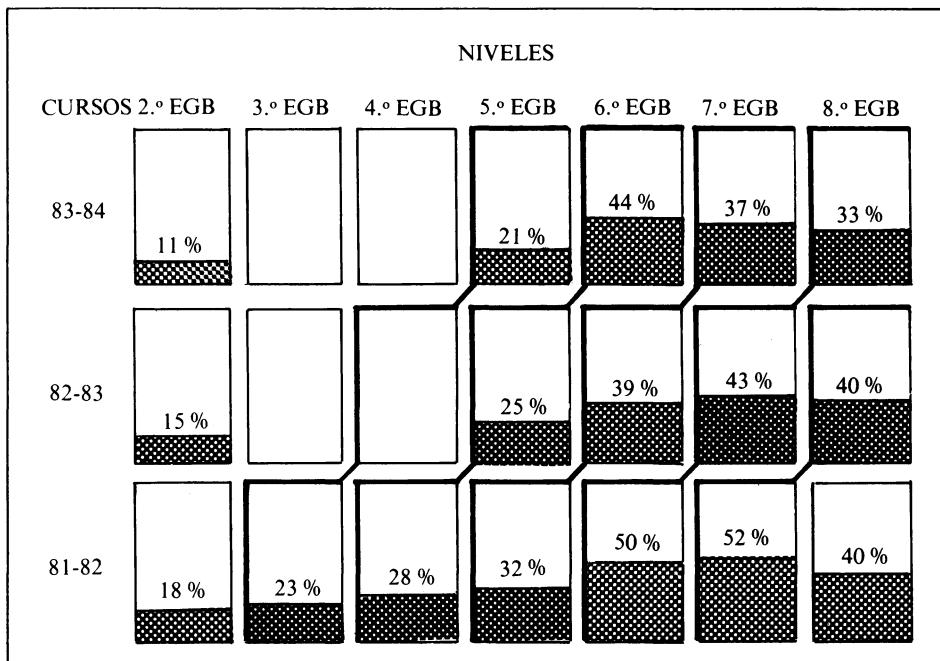


Gráfico 4: Porcentajes de insuficientes en el área de Matemáticas (N.B.: No se computan las tres últimas aulas puestas en funcionamiento).

La mayoría de estos alumnos proceden de familias con un nivel socioeconómico bajo y con estudios de escolarización básica, más bien deficientes.

Los casi 150 profesores que participan directamente en la experiencia llevan un promedio de docencia de 14 años los del C. Inicial, 16 los del Medio y 11 los del Superior. Los primeros tienen un promedio de permanencia en el centro de 5 años, los segundos de 6 años y los terceros de 4 años. Si se exceptúan los profesores de 8.º de EGB, son escasos los que permanecen con el mismo grupo más de dos cursos.

En cuanto a status profesional, la mayoría son propietarios definitivos, con solo un 10% de provisionales y un 5% de interinos, que se concentran en el Ciclo Superior.

### c) Aulas fijas control

Los 1.342 alumnos de este grupo fueron elegidos con idéntico sistema al anterior, buscando el máximo de similitudes con aquél y cumpliendo como condición indispensable el pertenecer al mismo barrio. En algún caso, como en Hospitalet, la mitad del centro asociado actúa como control y la otra mitad como experimental.

En resumen, los resultados que a continuación vamos a ofrecer, se han obtenido del número de alumnos que se detallan en el gráfico 5, para cada uno de los objetivos expuestos.

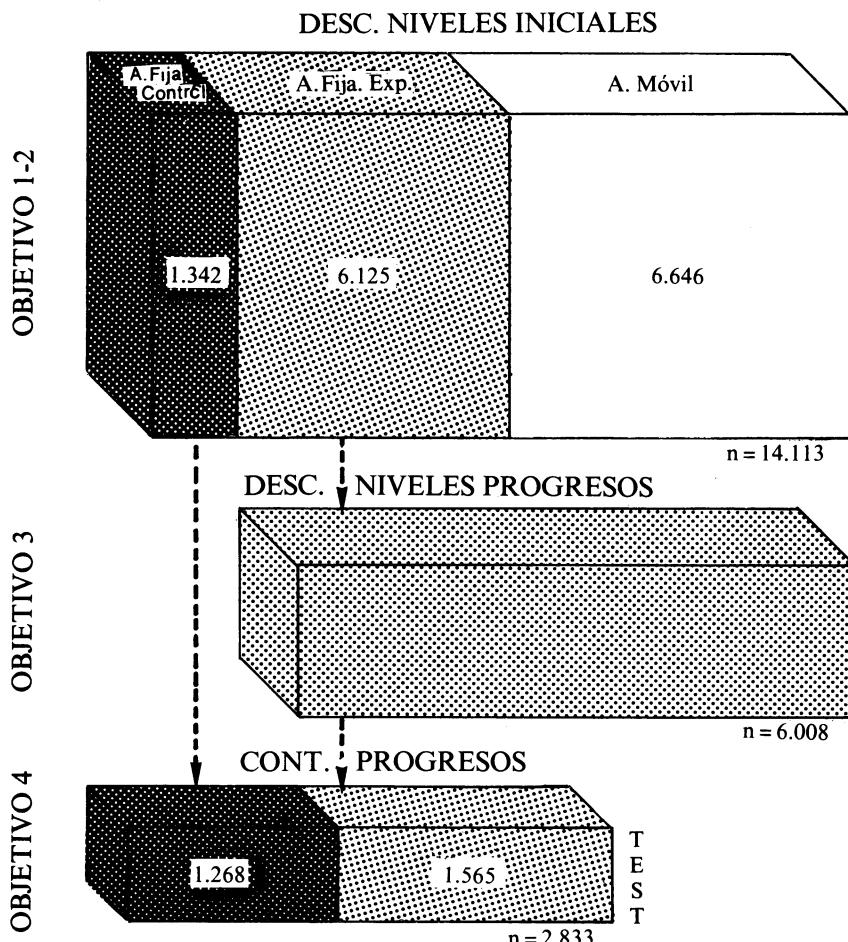


Gráfico 5: Número de alumnos en los que se ha contrastado cada uno de los objetivos de la investigación para el curso 84-85.

### OBJETIVO 1.<sup>o</sup>

A lo largo del curso 84-85, 14.113 alumnos han realizado el test del sistema, que, como ya dijimos, está pensado para identificar los niveles de aprendizaje en cada uno de los temas del «courseware».

Conocíamos las demostraciones del CET (Osin, L, 1981) sobre la consistencia del test, cuando los alumnos son inmediatamente «retestados». Teníamos presentes las correlaciones significativas entre niveles del test y puntuaciones del SAT (Stanford Achievement Test), detectados en el CCC (Macken, E, el alt. 1.976). A pesar de ello, consideramos prudente retomar el estudio de su consistencia, porque de ella podría depender una correcta prospección de conocimientos y el análisis de los progresos esperados en los alumnos del grupo experimental.

### **Consistencia de la prueba**

En nuestro estudio no hemos podido contar con criterios externos de medida independiente y se han utilizado estudios correlacionales entre las valoraciones del profesorado, previas a la ejecución del test, y los resultados del mismo, así como entre los resultados de éste (test) y una segunda ejecución (retest), realizada a continuación por los mismos alumnos.

Las correlaciones entre las valoraciones previas del profesorado y los niveles alcanzados por el alumno son casi todas significativas, pero no podemos asegurar que las primeras se hayan dado antes de conocerse los resultados de las doce sesiones del test, por lo que su validez es relativa.

El estudio correlacional entre test y retest, realizado en el C.P. «Montfalgars» (Girona) entre el 21-II-85 y 12-III-85, mostró en su totalidad relaciones significativas ( $\alpha = .05$ ), tanto entre las medias globales, como entre los temas «abiertos» al conjunto de cada uno de los cursos (Véase Tabla 1).

TABLA 1

Coeficientes de correlación entre test-retest  
(C.P. Montfalgars)

<b>CURSOS</b>	<b>Medias</b>	<b>Temas</b>				
		<b>1</b>	<b>8</b>	<b>9</b>	<b>14</b>	
2. <sup>º</sup> EGB	.49	.82	.81	.68	.67	
3. <sup>º</sup> EGB	.52	.87	.71	.82	.82	.81
4. <sup>º</sup> EGB	.58	.92	.85	.80	.88	.86
5. <sup>º</sup> EGB	.27	.88	.77	.77	.90	.93
6. <sup>º</sup> EGB	.19	.85	.73	.72	.79	.65
7. <sup>º</sup> EGB	.8	.95	.93	.95	.88	.96

## Resultados de la prospección

Desde que a principios de Noviembre de 1984 se conocen los primeros resultados del test (Véase Tabla 2), hasta obtener los últimos en el mes de Junio de 1.985 (Véase Tabla 3), han pasado ocho meses de curso en los que cabe esperar unos progresos (Gráfico 7), explicables por la acción docente ejercida y la normal ejercitación de los alumnos, al margen del sistema TOAM.

Teóricamente los niveles de aprendizajes que deberían mostrar estos alumnos en cada una de estas épocas habrían de ser los reseñados en ambas tablas en la fila de «Niveles ideales» y en el gráfico con la «Recta ideal».

**TABLA 2**  
Niveles de aprendizajes aritméticos reales e ideales (NOVIEMBRE 84)

A. FIJAS	E S t	CURSOS							
		2º	3º	4º	5º	6º	7º	8º	
ZONAS		$\bar{X}$	23.5	29.4	36.1	41.2	44.5	47.4	49.7
Barcelona		s	5.0	5.5	7.2	6.7	5.9	8.1	9.5
Terrassa		n	244	237	250	235	234	231	215
Hospitalet			23	33	43	53	63	73	
N. IDEALES									
A. MOVIL									
ZONAS		$\bar{X}$	20.7	28.9	34.8	39.5	42.1	45.2	51.4
Vic		s	4.0	6.5	6.7	6.7	5.9	6.8	9.5
		n	161	174	159	161	229	232	191
N. IDEALES			23	33	43	53	63	73	

**TABLA 3**  
Niveles de aprendizajes aritméticas reales e ideales (JUNIO 85)

A. MOVIL	E S T	CURSOS							
		2º	3º	4º	5º	6º	7º	8º	
ZONA		$\bar{X}$	29.0	33.5	40.0	43.1	44.8	47.9	54.1
S. HIPOLIT		s	6.6	7.5	6.8	6.7	6.9	9.0	11.9
		n	102	115	124	121	112	133	82
N. IDEALES			30	40	50	60	70		

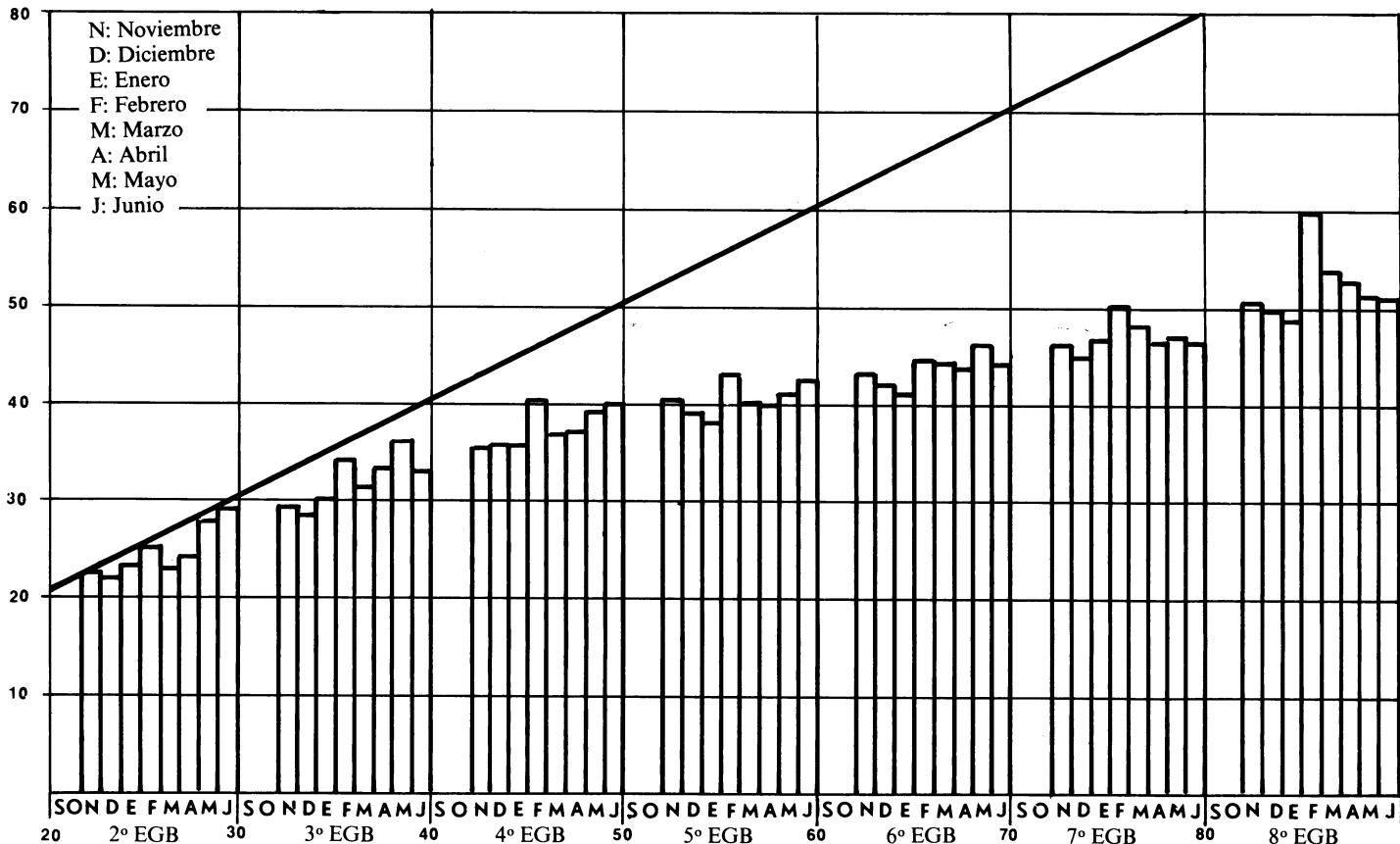


Gráfico 7: Evolución de niveles promedios en los ocho meses de ejecución del test.

El gráfico refleja una tendencia de niveles bastante uniforme que prácticamente sólo se rompe en el mes de febrero. Estos corresponden a la zona de Manlleu y por el momento no contamos con razones suficientemente contrastadas que expliquen esta separación con respecto a la norma.

Pero más allá de este hecho, se observan fenómenos de gran importancia como es la notable discrepancia entre los niveles empíricos y los criteriales (recta del gráfico 7), que se hace más patente conforme avanzan los cursos:

- \* Sólo los alumnos de 2.<sup>º</sup> de EBG mantienen en la primera mitad de curso concordancia entre ambos niveles. Después, éstas dejan de presentarse en la mayoría de los grupos analizados.
- \* En 3.<sup>º</sup> y 4.<sup>º</sup> de EGB los retrasos comienzan a perfilarse con más claridad y al final del Ciclo Medio (5.<sup>º</sup> de EGB) son ya más de un año.
- \* En el Ciclo Superior, los alumnos de 6.<sup>º</sup> muestran retrasos de aproximadamente dos años, los de 7.<sup>º</sup> de dos años y medio y los de 8.<sup>º</sup>, que teóricamente deberían superar el nivel máximo del «courseware» (79), mantienen en general niveles de dominio propios de los inicios de un 5.<sup>º</sup> de EGB.

Esta situación ha venido a corroborar una de las conclusiones del Informe del M.E.C. (1.984) sobre la *Evaluación de las Enseñanzas Mínimas, Ciclo Inicial de la EGB*. En este estudio se afirma que «la acumulación de fallos, lagunas y, en definitiva, la falta de dominio necesario para seguir aprendizajes posteriores,... pueden originar retrasos permanentes».

## OBJETIVO 2.<sup>º</sup>

Resulta curioso observar que este panorama últimamente descrito sobre pasa las fronteras nacionales, y, según la experiencia acumulada por el CET, se repite prácticamente en todos los entornos por ellos estudiados. Los resultados de sus investigaciones sitúan alrededor de cuatro niveles los progresos entre cursos, lo que supone un dominio medio de sólo el 40% de los requerimientos del programa de trabajo.

El dato era suficientemente grave como para darlo por válido sin más y además, necesitábamos contar con niveles iniciales de referencia contrastables con ese 100% de progreso, que los promotores del sistema auguraban, tras la fase de ejercitación con el mismo. Estos motivos nos llevaron a replicar el estudio en Catalunya y se ha utilizado el mismo sistema de cómputo empleado por ellos.

Este se basa en dividir la diferencia entre el curso más alto y más bajo, que participan en la experiencia, por el número de cursos que median entre ambos. En nuestro caso:  $X_p = \bar{X}_{8.\text{º} \text{EGB}} - \bar{X}_{2.\text{º} \text{EGB}}/6$

Como en el *Proyecto EAO* el test, que se utiliza como prueba de referencia, no lo han realizado este curso todos los alumnos al mismo tiempo, se ha tenido

en cuenta esta variable, tal y como se recoge en el Gráfico 8. Como se puede observar, el promedio de progresos aritméticos entre 2.<sup>º</sup> y 8.<sup>º</sup> de EGB es en la muestra analizada siete décimas superior al identificado por el CET.

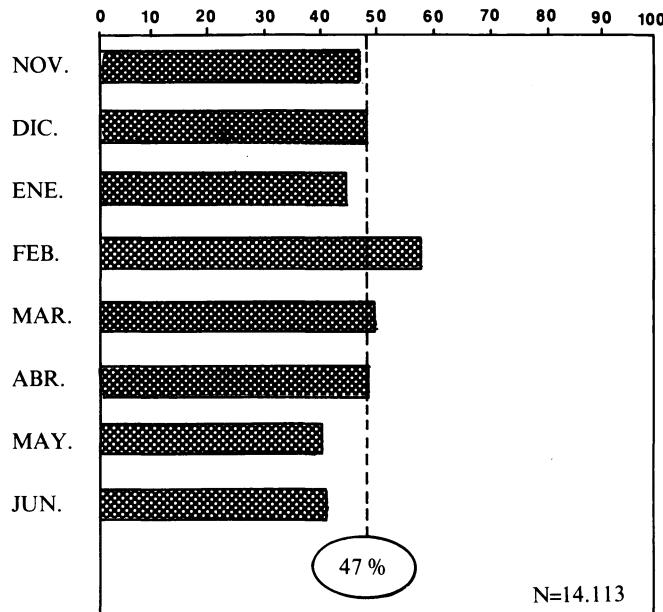


Gráfico 8: Porcentaje de progresos medios de niveles en función de los meses de realización del test.

A pesar de ello, el estado de la cuestión sigue siendo preocupante, puesto que sólo se cubre el 47% de los requerimientos del programa y además este porcentaje no se distribuye de manera uniforme entre los cursos, como se observa en el Gráfico 9

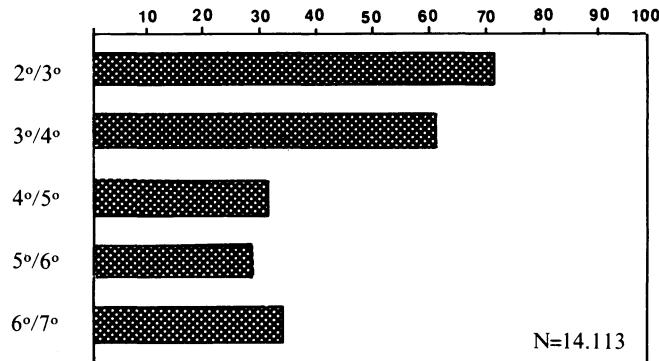


Gráfico 9: Porcentaje de progresos medios de niveles entre cursos.

Ante estos datos permítasenos dejar en el aire algunos interrogantes, que ya han sonado en la reunión de investigadores educativos del «Conseil de l'Europe» sobre la Renovación de la Enseñanza de las Matemáticas en la Educación Primaria (DECS, 1.984): ¿La acción docente se olvida del cálculo aritmético a partir de los primeros cursos de la educación primaria?, ¿Sigue aún en pie la falsa oposición entre las llamadas «antigua» y «moderna» matemáticas?.

### OBJETIVO 3.º

Acabamos de describir los niveles de aprendizaje y promedios de progreso que muestran los alumnos, antes de entrar en la fase de ejercitación. Queda por comprobar la incidencia de esta fase en el grupo de alumnos ( $n = 6.008$ ) que la han realizado.

De acuerdo a los promotores del sistema, si todos ellos hubiesen ejercitado durante los diez meses de curso, cabría esperar un progreso de diez niveles por año –100 % de progreso– o lo que es lo mismo un nivel por mes. Como la experiencia comenzó iniciado el curso y no todos ellos han trabajado el mismo tiempo, debemos esperar unos progresos medios similares a los reseñados en la última fila de la Tabla 4.

**TABLA 4**  
Esperanzas de progreso en función del tiempo  
que media entre test inicial y final

Meses entre tests:	7	6	5.5	5	4	3	2.5
Horas de ejercitac./ alumno entre tests	11,30	10	8	10,4	8	2,4	4,2
Número de alumnos	1.640	1.064	1.110	692	662	573	267
Esperanza niveles Progreso	7	6	5.5	5	4	3	2.5

Nota: El tiempo de realización de cada test es de 2 horas no computadas en esta Tabla.

Para comprobar la hipótesis planteada se han contrastado los resultados de los tests que se realizaron antes y después de la fase de ejercitación. La evolución de los mismos se puede observar en el Gráfico 10 con los perfiles medios que corresponden a cada curso, teniendo en cuenta los meses que han mediado entre ambas pruebas. Los principales estadísticos de las mismas se recogen en la Tabla 5.

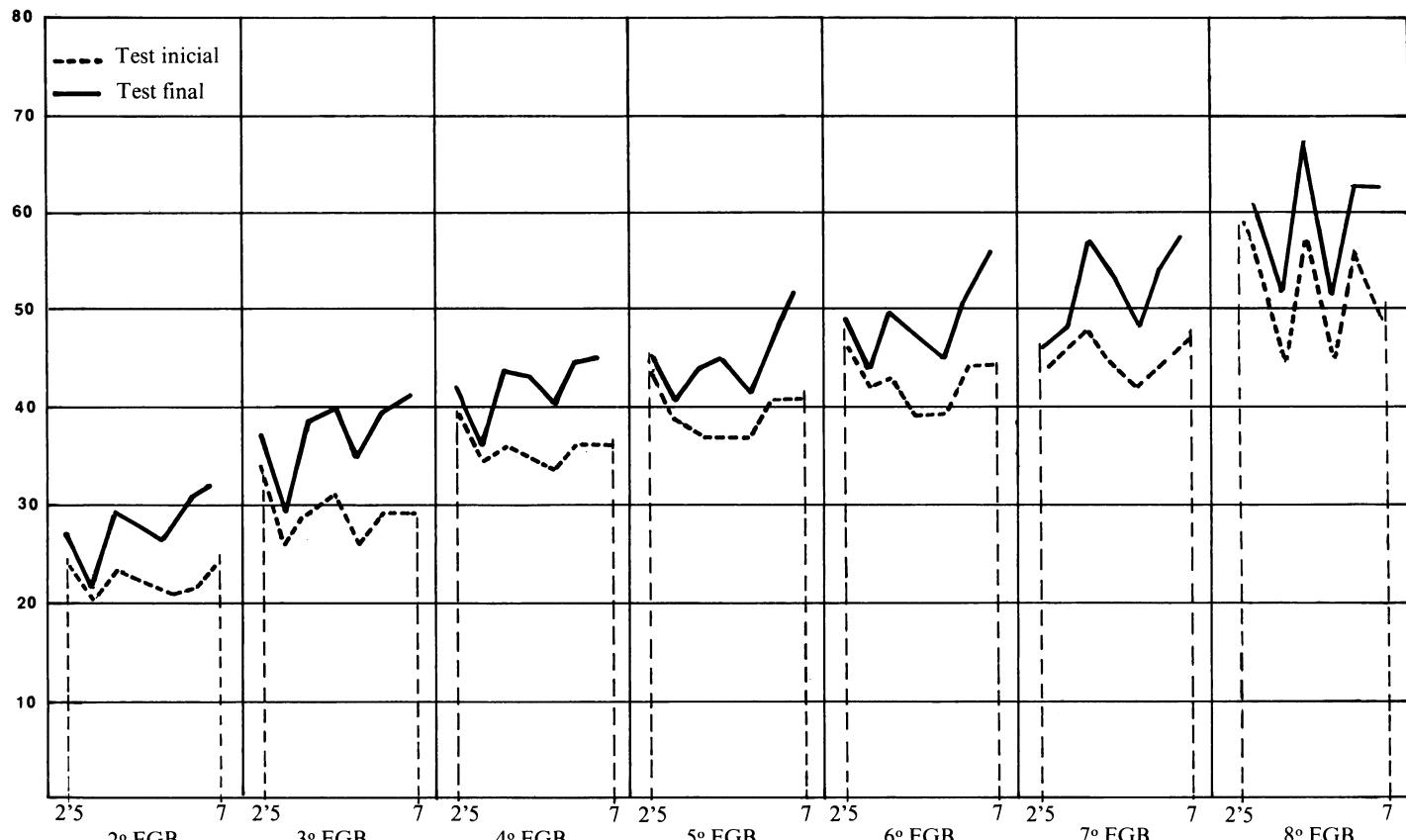


Gráfico 10: Evolución de los niveles promedios en función de meses que median entre el test inicial y final.

**TABLA 5**  
**Principales estadísticos del test inicial y final en función de los meses que median entre ambos**

		CURSOS									
		2º	3º	4º	5º	6º	7º	8º	Tot.		
<b>7</b>	Inic.	{ X s n	23.5 5.0 244	29.4 5.5 237	36.1 7.2 250	41.2 6.7 235	44.5 5.9 234	47.4 8.1 231	49.7 9.5 215		
										1.646	
	Final	{ X s n	32.0 7.0 243	41.0 6.1 235	45.5 7.5 249	52.0 9.7 236	56.1 10.3 233	58.2 12.3 230	62.0 12.0 214		1.640
-----		PROGRESOS		8.5	11.6	9.4	10.8	11.6	10.8	12.3	
<b>6</b>	Inic.	{ X s n	22.3 4.2 152	29.5 6.2 161	36.5 6.6 160	39.8 6.6 168	43.4 6.2 132	44.1 5.7 135	53.1 10.4 124		1.032
	Final	{ X s n	30.5 6.9 156	39.4 5.9 160	45.3 7.0 158	46.8 7.1 168	51.2 10.1 164	53.8 9.7 135	62.7 15.5 123		1.062
-----		PROGRESOS		8.2	9.9	8.8	7.0	7.8	9.7	9.6	
<b>5,5</b>	Inic.	{ X s n	20.9 4.0 202	26.1 5.9 207	33.7 7.0 169	37.5 7.2 203	39.4 7.0 137	42.4 7.5 135	45.6 9.0 107		1.160
	Final	{ X s n	26.6 6.3 195	34.6 7.2 193	40.4 7.4 161	41.7 8.9 200	44.9 8.7 135	48.1 8.8 127	51.7 9.2 99		1.110
-----		PROGRESOS		5.7	8.5	6.7	4.2	5.5	5.7	6.1	
<b>5</b>	Inic.	{ X s n	21.7 4.3 92	30.7 6.2 94	34.7 7.2 125	36.9 7.3 134	39.4 7.0 102	45.1 6.5 100	55.5 10.1 47		694
	Final	{ X s n	28.2 5.5 93	40.0 6.9 94	43.2 7.4 123	45.0 8.7 134	47.4 9.5 101	53.2 10.1 100	67.4 11.7 47		692
-----		PROGRESOS		6.5	9.3	8.5	8.1	8.0	8.1	11.9	

			CURSOS								tot.
			2. <sup>º</sup>	3. <sup>º</sup>	4. <sup>º</sup>	5. <sup>º</sup>	6. <sup>º</sup>	7. <sup>º</sup>	8. <sup>º</sup>		
<b>4</b>	M	Inic.	X	23.3	30.3	36.3	37.5	42.9	48.0	45.0	660
			s	4.4	6.4	6.0	7.0	7.7	8.7	9.4	
			n	78	99	99	92	92	98	102	
<b>S</b>	E	Final	X	29.5	39.0	44.0	44.2	50.1	57.0	51.9	662
			s	7.0	6.4	5.4	7.1	9.5	11.0	12.0	
			n	79	99	99	93	92	98	102	
-----			PROGRESOS	6.2	8.7	7.7	6.7	7.2	9.0	6.9	
<b>3</b>	M	Inic.	X	20.5	25.8	34.8	39.0	42.3	46.2	55.4	556
			s	5.3	7.2	8.4	8.2	7.0	7.0	8.3	
			n	94	75	91	110	87	52	47	
<b>S</b>	E	Final	X	21.4	28.6	35.9	40.7	43.7	47.7	59.8	573
			s	6.0	9.1	9.8	8.4	10.2	8.9	10.9	
			n	99	75	94	116	90	52	47	
-----			PROGRESOS	0.9	2.8	1.1	1.7	1.4	1.5	4.4	
<b>2,5</b>	M	Inic.	X	23.7	33.8	39.4	43.6	46.5	43.6		267
			s	5.4	6.1	7.8	5.3	5.4	10.1		
			n	71	69	65	33	21	8		
<b>S</b>	E	Final	X	27.1	37.1	42.3	44.8	48.7	46.1		267
			s	6.9	7.4	7.6	6.0	6.9	7.6		
			n	70	69	65	33	21	9		
-----			PROGRESOS	3.4	3.3	2.9	1.2	2.2	2.5		

Una simple comparación entre los niveles de progresos alcanzados (Tabla 5) y los esperados para cada uno de los meses que median entre el test inicial y final (última fila Tabla 4), evidencia que los primeros han superado a los segundos en todos los casos excepto en las clases de 2.<sup>º</sup> a 7.<sup>º</sup> del centro de Acción Especial «St. M.<sup>a</sup>. Gardeny» (3 meses entre ambos), en 5.<sup>º</sup> de EGB con un intermedio de cinco meses y medios y en 5.<sup>º</sup> y 6.<sup>º</sup> con dos meses y medio.

Tras realizar los contrastes pertinentes de medias, se puede concluir que prácticamente en todos los casos se observan diferencias significativas al final de curso, que son más ostensibles conforme aumenta el número de horas dedicado a la ejercitación.

En este contexto cabe resaltar que los 1.640 alumnos que han realizado aproximadamente 11,30 horas/ curso de ejercitación han conseguido un progreso medio de once niveles frente a los siete esperados, lo que supone una supera-

ción media del 153% de los requerimientos del programa, que han repercutido en que los niños que en el curso 84-85 cursaban:

- \* 2.<sup>º</sup> y 3.<sup>º</sup> de EGB, emprenderán los cursos siguientes con niveles de aprendizaje superiores a los niveles ideales.
- \* 4.<sup>º</sup> y 5.<sup>º</sup> de EGB, comenzarán sus nuevos cursos con una reducción de discrepancias sobre la tendencia general de cinco y ocho niveles respectivamente.
- \* 6.<sup>º</sup> de EGB, ya no iniciarán séptimo con dos cursos de retraso, sino con curso y medio.
- \* 7.<sup>º</sup> y 8.<sup>º</sup>, probablemente terminarán la EGB o iniciarán las Enseñanzas Medias con serios déficits aritméticos, pero no tan graves como al inicio de la experiencia.

Similar tendencia se observa en los alumnos del C. Inicial y Medio con seis meses entre pruebas, aunque los retrasos, siguen siendo graves en el C. Superior. En este grupo se ha conseguido un promedio de progreso de *nueve niveles frente a los seis esperados*.

Conforme disminuye el número de horas de ejercitación, la incidencia del sistema no es tan significativa, pero se percibe su influencia.

#### OBJETIVO 4.<sup>º</sup>

Para un mayor control de la incidencia del *Proyecto*, se ha llevado a término el diseño cuasi-experimental, reflejado en el Gráfico 11.

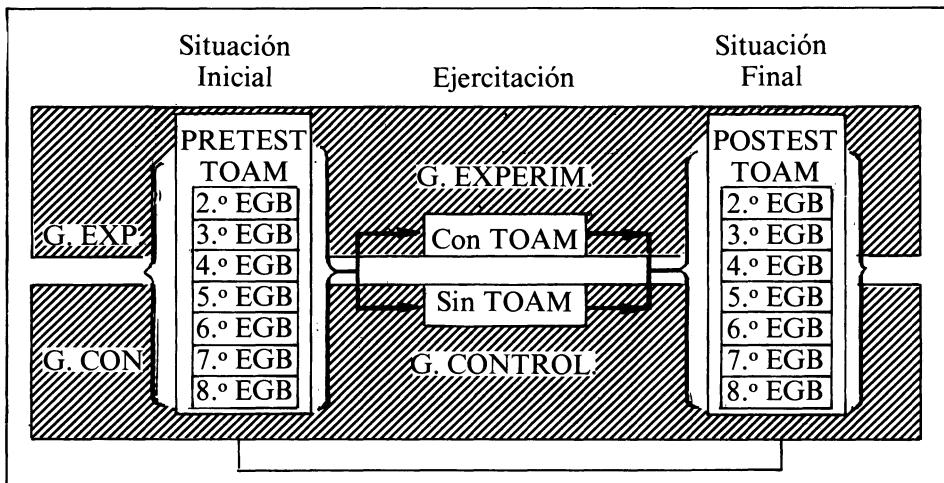


Gráfico 11: Diseño cuasi-experimental de los resultados.

Los alumnos de dos centros públicos de Barcelona y Terrassa, situados en el mismo barrio en el que se ubican aulas fijas experimentales, actúan como grupo control. Y en Hospitalet, la mitad de los alumnos de los centros «Milagros Consarnau» y «Bernat Desclot» actúan como grupo control y la otra mitad como experimental.

Por cuestiones de organización ni todos los grupos experimentales han dispuesto de las mismas horas de ejercitación (Véase segundo columna Tabla 8), ni el período que medió entre el pretest y el postest fue similar para los grupos control y experimental (Véase primera columna Tabla 8).

### Situación inicial

Los resultados del pretest se recogen en la Tabla 6, diferenciados por ciudades y en el caso de Hospitalet por centros.

Al no poder elegir grupos homogéneos por razones deontológicas, nos encontramos en esta situación inicial con diferencias significativas de medias a favor del grupo experimental en los cursos 2.<sup>º</sup>, 4.<sup>º</sup>, 5.<sup>º</sup>, 6.<sup>º</sup> y 7.<sup>º</sup> de Barcelona y 3.<sup>º</sup>, 5.<sup>º</sup>, 6.<sup>º</sup>, 7.<sup>º</sup> y 8.<sup>º</sup> de Terrasssa. A su vez, la diferencia de medias es significativa a favor del grupo control en 5.<sup>º</sup> del «M. Consarnau» y 3.<sup>º</sup> del «B. Desclot».

### Situación final

Los resultados del postest se recogen en la Tabla 7, de idéntica forma a como se hizo en la situación inicial.

Los contrastes de medias pretest-postest arrojan en su mayoría diferencias significativas ( $\alpha = .05$ ) a favor de la situación final tanto en los grupos control como experimentales, como cabía esperar por la labor normalmente realizada a lo largo de un curso con o sin ayuda tecnológica. Sin embargo, los niveles de progreso son marcadamente diferentes para ambos grupos, como se puede observar en la Tabla 8.

Si se exceptúa el grupo de 8.<sup>º</sup> de EGB del «M. Consarnau», en todas las demás clases los mayores niveles de progreso se dan en los grupos experimentales. Prácticamente en todos ellos, menos en 5.<sup>º</sup> y 8.<sup>º</sup> del «M. Consarnau» y 6.<sup>º</sup> del «B. Desclot», se superan los niveles esperados de progreso, mientras que en los del control todas las clases quedan por debajo de los mismos con la excepción de 2.<sup>º</sup> de EGB del «M. Consarnau».

Esta situación ventajosa de los grupos experimentales se ve más clara a medida que aumenta el número de horas de ejercitación y de meses entre ambas pruebas. Así los grupos de Barcelona y Terrassa logran superar en un 170% la esperanza de progreso y sus grupos parejos de control sólo alcanzan el 57% y 70% respectivamente.

No obstante, en la lectura de todos estos resultados hay que tener en cuenta

TABLA 6  
Principales estadísticos del Pretest  
en grupos experimentales y de control

		CURSOS							Tot.	
		2º	3º	4º	5º	6º	7º	8º		
B A R C	G. EXP.	$\bar{X}$	25.9	29.9	35.9	44.5	45.4	49.3	49.2	474
		s	5.8	5.7	6.9	6.8	5.7	10.3	11.5	
		n	66	69	69	68	69	70	63	
E L O N	G. CON.	$\bar{X}$	21.7	29.3	37.6	38.9	41.7	44.5	46.9	442
		s	4.0	6.9	7.0	6.4	7.8	7.0	11.0	
		n	63	65	56	62	69	66	61	
A	Dif. sig.:		&	&	&	&	&	&		
T E R R A S A	G. EXP.	$\bar{X}$	22.0	28.8	35.9	39.4	44.1	47.7	50.5	728
		s	3.6	5.1	7.0	6.5	4.8	6.6	7.1	
		n	113	108	118	108	98	94	89	
R R A S A	G. CON.	$\bar{X}$	21.9	27.0	35.6	36.5	41.2	43.3	45.9	462
		s	3.5	5.5	7.5	7.7	5.7	6.1	7.8	
		n	64	64	68	68	67	67	64	
A	Dif. sig.:		&	&	&	&	&	&		
M C O N S A R N	G. EXP.	$\bar{X}$	23.3	33.1	36.4	36.8	42.1	42.3	47.3	220
		s	5.0	5.6	6.9	8.7	7.3	6.2	8.9	
		n	31	29	30	33	31	33	33	
S A R N	G. CON.	$\bar{X}$	23.2	33.1	37.1	42.7	39.1	42.8	46.5	222
		s	5.0	5.6	5.8	4.3	6.4	4.9	8.0	
		n	32	30	33	33	32	32	30	
N	Dif. sig.:									
B D E S	G. EXP.	$\bar{X}$	24.4	26.2	38.0	38.8	41.0			146
C		3.8	7.0	6.4	5.9	6.4				
L O T		n	27	30	29	29	31			
C L O T	G. CON.	$\bar{X}$	22.8	29.2	35.9	40.6	43.6			151
		s	4.6	4.6	6.3	9.2	5.0			
		n	27	32	30	31	31			
	DIF. SIG.									

(&) = Diferencia significativa ( $\alpha = .05$ ) entre medias

**TABLA 7**  
**Principales estadísticos del postest**  
**en grupos experimentales y de control**

		CURSOS								Tot.
		2º	3º	4º	5º	6º	7º	8º		
B	G. EXP.	$\bar{X}$	32.7	41.9	45.9	59.6	58.1	65.0	59.9	471
		s	6.2	6.5	7.3	9.5	9.0	12.4	14.0	
		n	66	68	69	68	68	69	63	
C										
E	G. CON.	$\bar{X}$	26.7	33.8	41.2	40.7	44.3	47.5	51.1	438
L		s	6.2	7.1	7.0	7.0	8.0	8.8	12.3	
O		n	59	65	56	62	69	66	61	
N										
A	Dif. sig.:		&	&	&	&	&	&		
T	G. EXP.	$\bar{X}$	32.6	40.6	44.5	49.0	54.1	56.8	62.9	728
		s	6.8	5.4	6.7	7.6	10.2	11.3	10.3	
		n	113	108	118	108	98	94	89	
R										
R	G. CON.	$\bar{X}$	25.0	32.6	39.8	39.5	43.2	46.2	49.4	462
A		s	4.4	6.9	7.0	7.4	5.6	6.8	9.7	
S		n	64	64	68	68	67	67	64	
A	Dif. sig.:		&	&	&	&	&	&		
M.	G. EXP.	$\bar{X}$	33.7	42.4	43.9	46.0	46.6	48.8	47.4	222
C		s	5.7	4.1	7.9	10.0	8.5	7.7	14.4	
O		n	31	30	30	33	32	33	33	
N										
S	G. CON.	$\bar{X}$	29.5	36.4	40.0	46.2	41.4	46.2	46.6	219
A		s	7.3	7.1	6.4	4.5	6.6	4.7	8.9	
R		n	32	29	32	33	32	32	29	
N	Dif. sig.:		&	&	&					
B.	G. EXP.	$\bar{X}$	33.7	36.1	46.1	45.7	46.3			144
D		s	5.6	7.2	7.0	6.5	8.4			
E		n	28	29	27	29	31			
S										
C	G. CON.	$\bar{X}$	27.1	35.1	40.2	45.4	47.7			149
L		s	6.5	5.8	6.8	10.6	7.7			
O		n	27	31	30	31	30			
T	DIF. SIG.		&		&					

**TABLA 8**  
**Niveles de progresos medios entre pretest  
y postests en grupos experimentales y control**

	Mes. int. Test	Horas Ejerc Toam	Cursos								Nivel Gral.	progr Porcn.
			2. <sup>º</sup>	3. <sup>º</sup>	4. <sup>º</sup>	5. <sup>º</sup>	6. <sup>º</sup>	7. <sup>º</sup>	8. <sup>º</sup>			
BARC. { G. EX. G. CO.	7	13.2	6.8	12.0	10.0	15.1	12.7	15.7	10.7	11.9	170%	57%
	6	—	5.0	4.5	3.6	1.8	2.6	2.0	4.7	3.4	3.4	
TERR. { G. EX. G. CO.	7	10.4	10.6	11.8	8.6	9.6	10.0	9.1	12.4	10.3	171%	70%
	5	—	3.1	5.6	4.2	3.0	2.0	2.9	3.5	3.5	3.5	
M. CO. { G. EX. G. CO.	6	10.4	10.4	9.3	7.5	9.2	4.5	6.5	0.1	6.8	113%	57%
	6	—	6.3	3.3	2.9	3.5	2.3	3.4	0.1	3.4	3.4	
B. DE. { G. EX. G. CO	6	8	9.3	9.9	8.1	6.9	5.3			7.9	132%	78%
	6	—	4.3	5.9	4.3	4.8	4.1			4.7	4.7	

la deficiente homogeneidad de la situación de partida, cuyas diferencias han sido siempre favorables a los grupos experimentales.

Esta circunstancia se ha paliado en los otros dos centros, cuyos grupos control y experimental pertenecen al mismo colegio y el tiempo transcurrido entre el pretest y postest es el mismo. En este caso los alumnos que actuaban como grupo experimental han superado la esperanza de progreso en un 113% y 132%, mientras que sus compañeros del control lo han hecho en un 57% y 78% respectivamente.

Por la forma en que se ha llevado a término este último análisis, puede resultar prematuro asegurar que se ha comprobado la eficacia del *Proyecto EAO* en este su primer año de experiencia y que los progresos observados en los grupos experimentales se deban únicamente a la introducción del soporte informático. Lo que si es evidente una serie de hechos que antes de ponerse en práctica no se daban, y al final de curso se producen:

- Todos los grupos experimentales de 2.<sup>º</sup> de EGB parten hacia tercero con un promedio de niveles por encima de los requeridos por el programa, hecho que no se presenta en ninguno de los grupos control.
- De similar situación gozarán todos los grupos experimentales de 3.<sup>º</sup> de EGB que pasen a cuarto, menos los del «B. Desclo». Por el contrario todos los grupos control empezarán el curso siguiente por debajo de los niveles requeridos.
- A partir de 4.<sup>º</sup> de EGB, los demás cursos de los grupos experimentales han reducido sus iniciales discrepancias entre niveles ideales y reales, mientras que en los grupos control vuelve a repetirse similar tendencia a la mostrada en la prospección de niveles de aprendizaje.

## CONCLUSIONES

Existe una tendencia de glorificar las nuevas tecnologías de la información y de casi convertirlas en la panacea que resuelve todos los problemas de la enseñanza-aprendizaje por si solas. Antes ocurrió con las otras «nuevas» tecnologías y el riesgo no ha desaparecido.

Quisieramos huir de este vicio y simplemente reconocer que «el empleo de los ordenadores en la educación... es un fenómeno imparable». De lo que se trata es de obtener de ellos el mejor rendimiento posible, de tal forma «que se conviertan en un elemento innovador y no que sirvan para vestir con mismos ropajes las mismas prácticas escolares que se realizan hace siglos y que no satisfacen a las necesidades de los escolares» (Del Val, J. 1.985).

Con este espíritu asumimos la investigación y al poco tiempo de trabajar en ella empezamos a compartirlo con un buen grupo de docentes, miembros activos en el Proyecto. Ellos, más que nadie, viven de cerca la grave situación, detectada en la prospección de niveles de aprendizajes aritméticos con alumnos del:

- *Ciclo Inicial* que pronto quedan rezagados en los rudimentos básicos del cálculo aritmético.
- *Ciclo Medio* que muestran al final del mismo un retraso de un curso escolar con respecto a los niveles requeridos.
- *Ciclo Superior* con un promedio de superación de cálculos aritméticos que corresponden por currículum a un 5.<sup>º</sup> de EGB, porque en 7.<sup>º</sup> llevaban casi dos cursos y medio de retraso en 6.<sup>º</sup> de dos años.

En casi todos los centros donde estos niños estudian, se disponen de recursos tecnológicos, más o menos «nuevos», que cuando aparecieron se mal presentaron como el remedio total a los problemas de aprendizaje. Por lo que se ve, esto no ha sido así y por prudencia cabe pensar que el computador por si solo tampoco va a hacer milagros.

Tras los resultados de este primer año de experiencia, que hemos venido comentando, existe la acción de unos docentes que desde los primeros momentos se han percatado de los servicios ofrecidos por el *Proyecto EAO*, pero que pronto han descubierto que el sistema con sus detenidos y continuados diagnósticos les estaba exigiendo actividades compensatorias, cambios metodológicos,... etc, que en el desempeño anterior de su función no se presentaban tan urgentes, quizás por desconocimiento de la realidad de sus alumnos.

A lo largo de esta exposición no hemos hablado de este fenómeno, que en la actualidad está en fase de estudio, pero que es un hecho que hemos vivido de cerca, sobre todo en aquellos centros que llevan más tiempo en el Proyecto.

Su acción, a lo largo de siete meses de trabajo, junto a los efectos derivados del sistema han hecho posible que alumnos de:

- *Segundo y tercero de EGB pasen al curso siguiente con niveles acordes a los requerimientos curriculares, hecho que al inicio de la experiencia no se presentaba, sobre todo en el último de ellos.*
- *Del resto del Ciclo Medio hayan aminorado sus discrepancias con respecto a los niveles ideales de aprendizaje, de tal forma que puedan iniciar 5.<sup>º</sup> de EGB con cinco niveles de retraso y 6.<sup>º</sup> con ocho.*
- *Ciclo Superior también hayan rebajado sus discrepancias entre niveles ideales y reales, aunque es poco probable que superen los requisitos curriculares de Aritmética dentro de su escolarización en la EGB.*

Es en este tipo de centros donde mejor se percibe la incidencia del *Proyecto EAO*, pero no ha sido menor en el resto de centros, si se pone en relación el tiempo de participación en el mismo con los resultados alcanzados. Prácticamente en todos los centros experimentales se han superado los niveles de progresos esperados y en algunos de ellos ostensiblemente, circunstancia ésta que contrasta con los progresos de los grupos control, que se mantienen entre la mitad y tres cuartos por debajo de los progresos esperados para los alumnos que se ejercitan con el sistema.

Es pronto para evaluar todo el alcance el *Proyecto EAO*, pero estos primeros resultados apuntan una serie de efectos positivos, para compensar los déficits aritméticos que sufren los alumnos del tipo de población estudiada. Va a ser necesario, por tanto, seguir de cerca el Proyecto con investigaciones más pormenorizadas y con ese espíritu comienza a gestarse en estos momentos lo que en un futuro puede llegar a ser una línea de investigación.

## PROSPECTIVA

Las necesidades descubiertas a lo largo del curso 84-85 nos han ido fijando próximos objetivos de investigación o de desarrollo, que en algunos casos ya están en fase de estudio. Así para el curso 85-86 se pretende:

- a) Realizar un análisis más profundo de la adecuación entre el sistema TOAM y las «Orientacions i Programes» del Departament d'Ensenyament y sobre el índice de dificultad que presentan los ejercicios (items) del «courseware» a la muestra catalana. A su vez se piensa relacionar este estudio con datos similares de otros países donde se utiliza el mismo sistema.
- b) Analizar las principales variables *psicopedagógicas, matemáticas e informáticas* implicadas en el Proyecto.
- c) Controlar los efectos del Proyecto en *discentes, docentes y entidades escolares*.
- d) Colaborar en la elaboración de nuevos programas de EAO, de acuerdo a las necesidades descubiertas.

Serán éstos probablemente los primeros trabajos con los que se iniciará la línea de investigación, que cuenta ya con un círculo de trabajo en torno al Gabinete de Ordenación Educativa de E. Primaria. Este ha puesto su fondo documental al servicio de las personas, que presentan anteproyectos de investigación razonables y viables, con el único compromiso de retornar al círculo una copia de la misma. A esta posibilidad ya se han unido alumnos de la especialidad de Pedagogía con vistas a elaborar sus Memorias de Licenciaturas y profesores ligados o no a la experiencia, pero que están interesados por el tema.

Para conjuntar todos estos esfuerzos se ha buscado un marco teórico que tuviese en cuenta los principales centros de interés que pueden acompañar al Proyecto y en principio se ha optado por un modelo interrelacionado similar al propuesto por Centra, J.A y Potter, D, A, (1980), pero sin perder de vista el efecto de «feedback» que los «resultados de aprendizaje» han de tener sobre el resto de «conductas», «características», «actividades» y «condiciones». (Véase Gráfico 12).

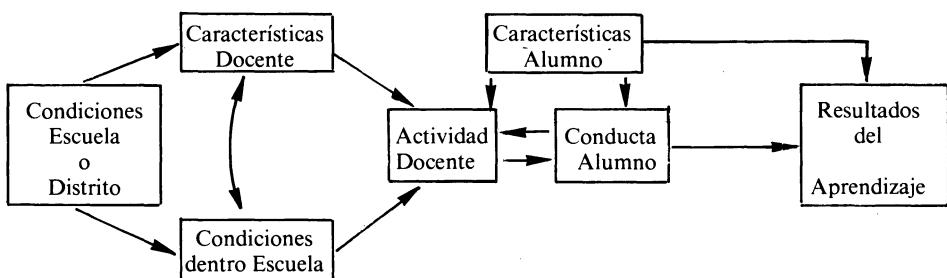


Gráfico 12: Modelo interraccional de análisis del rendimiento (Centra J.A. y Potter, D, A.).

Tal y como proponíamos en el II Seminario sobre Modelos de Investigación Educativa» (Echeverría, B. 1.983), en los primeros momentos se ha seguido una metodología de trabajo eminentemente descriptiva, a la vez que se han iniciado algunas investigaciones piloto de carácter explicativo y predictivo, centradas en las variables psicopedagógicas relacionadas con los resultados del aprendizaje. Sobre ellas se va a insistir, para facilitar la intervención docente que se presenta urgente en cuanto a didáctica de las Matemáticas se refiere.

Por otra parte, van a ser necesarias investigaciones longitudinales, cercanas a la línea desarrollada por Ragosta, M (1983), que permitan comprobar la eficacia del *Proyecto EAO* a lo largo de varios años y desvelar la existencia o no del efecto Hawthorne, tan ligado a veces a las nuevas tecnologías educativas.

Cabe esperar que, si el proyecto de investigación continúa con el necesario apoyo institucional, vayan apareciendo nuevos interrogantes a responder, como por ejemplo el estilo cognitivo o estrategias que utilizan los usuarios del Proyecto o los hábitos que se desarrollan en concomitancia con las estructuras motivacionales de docentes, discentes y entidades escolares.

## REFERENCIAS BIBLIOGRÁFICAS

- CENTRA, J. A. y POTTER, D. A. (1980).: «School and teachers effects: an interrelational model». *Rev. of Educational Research*, 50, 273-291.
- DESC/Rech (1984).: *Educational Research Workshop on the Renewal of Mathematics Teaching in Primary Education*. Conseil de l'Europe (Doc. policop.), Puidoux-Chexbres (Switzerland) 1-4 May 1984.
- ECHEVERRIA, B. (1983).: «La investigación empírica de carácter educativo en las universidades españolas (1980-1983)». *Revista de Investigación Educativa*, Vol. 1, N.º 2, 144-204.
- LORD, F. M. (1970).: «Some test theory for tailored testing» en HOLTZ MANN, W. H. (Ed).: *Computer-Assisted Instruction. Testing and Guidance*. Harper & Row, New York, 139-183.
- LORD, F.M. (1980).: *Applications of item response theory to practical testing problems*. Lawrence Erlbaum, Associates, Pub. New York, 150-160.
- MACKEN, E. y SUPPES, P. (1976).: *Evaluation Studies of CCC Elementary-School Curriculum: 1971-1975*. Computer Curriculum Corporation, Palo Alto.
- M.E.C. (1984).: *Evaluación de Enseñanzas Mínimas. Ciclo Inicial de la EGB*. Dirección Gral. de Educación Básica, Madrid.
- MINKOVICH, A., DAVIS, D. y BASHI, J. (1977).: *An evaluation of Israeli Elementary Schools*. The Hebrew University Of Jerusalem, Jerusalem.
- NESHER, P. (1976).: «Three determinants of difficulty in verbal arithmetic problems». *Educational Studies in Mathematics*, 7, 369-388.
- NESHER, P. y KATRIEL, T. (1977).: «A semantic analysis of addition and subtraction word problems in arithmetic». *Educational Studies in Mathematics*, 8, 251-269.
- NESHER, P. y KATRIEL, T. (1978).: «Two cognitive modes in arithmetic word problem solving». En *2nd Annual Meeting of the IGPME Osnabrück*, 226-241.
- NESHER, P. (1980).: «The stereotyped nature of school word problems» *For the Learning of Mathematics*, 1, 41-48..
- NESHER, P., GREENO, J. G y RILEY, M. S. (1982).: «The development of semantic categories of addition and subtraction». *Educational Studies in Mathematics*, 13, 373-394.
- OSIN, L. (1981) «Computer-Assisted instruction in arithmetic in Israeli disadvantaged elementary schools». En LEWIS, R & TAGG, D. (Ed) *Computers in education*. North-Holland, Pub., IFIP, 469-475.
- OSIN, L. (1984): *TOAM: CAI on a national scale*. IEEE, 418-424.
- RAGOSTA, M. (1983).: Computer-Assisted Instruction and compensatory Education». *Machine-Mediated Learning*, Vol 1, N.º 1, 97-127.
- SUPPES, P. y MORNINGSTAR. M. (1972): *Computer Assisted Instruction at Stanford (1966-1968)*. Academic Press, New York.
- DEL VAL, J. (1985): «Para qué vale un ordenador en la escuela». *Cuadernos de Educación y Nuevas Tecnologías de la Información*. 1. 4-9.

## «Tendencias actuales de la documentación científica automatizada en la investigación educativa

*por  
Luis Lizasoain*

### CUESTIONES GENERALES SOBRE DOCUMENTACIÓN CIENTÍFICA.

El actual desarrollo y evolución de las ciencias en general y de las llamadas sociales y humanas en particular conlleva una gran producción de documentación científica, que por su propia magnitud demanda la utilización de sistemas que permitan una eficaz y rápida recuperación y difusión de la misma para evitar la paradoja que pueda suponer el que dicho incremento de la información lleve implícito una cada vez mayor desconocimiento de los trabajos realizados en una disciplina cualquiera.

Se hace necesario pues disponer de sistemas que permitan el proceso de grandes masas de información y que además arbitren los canales de acceso y difusión para distribuirla allá donde se produce y necesita, evitando así la reduplicación de esfuerzos y facilitando la actualización de los conocimientos.

Evidentemente estas tareas son poco menos que irrealizables si no se cuenta con los recursos que proporcionan las nuevas tecnologías, especialmente la informática y la telemática.

En el ámbito de las Ciencias de la Educación también observamos como estas tecnologías están incidiendo prácticamente en todos los ámbitos de intervención educativa (proceso de enseñanza, aprendizaje, organización, administración evaluación y medición, diagnóstico...) y en concreto en el de la investigación nos encontramos con importantes aplicaciones en lo que se refiere a análisis de datos, modelos de simulación, muestreo, etc.

En esta línea de potenciación de la infraestructura de la investigación educativa, la documentación automatizada ofrece una serie de ventajas que no poseen los sistemas tradicionales. Pero antes de pasar a examinar más detenidamente estas cuestiones veamos brevemente algunos conceptos básicos de dicha disciplina con el fin de aclarar y fijar la terminología usada.

Cuando se habla de un sistema documental automatizado (por ejemplo ERIC, uno de los más difundidos en educación) son varios los elementos que lo

integran, pero simplificando mucho el modelo, dos se nos aparecen como fundamentales: de una parte la base de datos propiamente dicha y de otra el Tesauro de la misma.

Obviamos aquí elementos como la red de ordenadores y terminales, el logístico necesario para el tratamiento y recuperación de la información, y todo lo referente a los distribuidores y comercializadores de la base de datos. Elementos imprescindibles pero para nuestros fines tienen en este momento menos relevancia.

Comenzando por la base de datos propiamente dicha, en primer lugar aclararemos que usamos el término desde el punto de vista documental, pues en informática tiene otras acepciones y matices. En este caso lo podríamos definir como *el conjunto de referencias documentales almacenadas y organizadas sistemáticamente para recuperación y difusión selectiva en función de los criterios del usuario*.

En cuanto a los tipos de bases de datos se suele hablar, entre otras, de bases de datos bibliográficas, textuales o factuales. Ejemplos de cada una de ellas serían la ya citada de ERIC, una de sentencias jurídicas o de legislación y, por último, las cotizaciones de bolsa de un periodo cualquiera.

En nuestro ámbito, si exceptuamos lo referente a la legislación educativa, las más usuales y útiles son las bibliográficas en las que las fuentes primarias son artículos, conferencias, ponencias, informes, memorias de investigación, tesis y en ocasiones libros. La información se almacena según el modelo bibliográfico por diferentes campos (autor, título, identificadores, descriptores).

El general una base de datos implica, disponer, además de la infraestructura tecnológica antedicha, de los fondos documentales a vaciar, de un equipo de documentalistas que realicen dicha labor y, por último de un lenguaje controlado y normalizado que permita la «indización» del contenido documental de forma unívoca, de manera que los términos que utilice el documentalista a la hora de definir los conceptos que aparecen en el documento sean los mismos que emplee el usuario cuando quiera recuperar la información.

Este lenguaje controlado y normalizado se plasma en el Thesaurus o Tesauro (que sobre su denominación no hay excesivo consenso).

La norma ISO 2788 define un Tesauro como sigue:

- según su función, es un instrumento de control terminológico usado por los «indizadores» o usuarios para traducir el lenguaje natural de los documentos a un lenguaje más controlado.
- según su estructura, es un vocabulario dinámico y controlado de términos relacionados semántica y genericamente que cubre un dominio específico del conocimiento.

Un Tesauro se compone de *descriptores* (términos o expresiones que designan, unívocamente los conceptos del campo cubierto por el tesauro), *no descriptores* (sinónimos o quasi-sinónimos de los mismos) y *relaciones semánticas* (entre descriptores y no descriptores o entre descriptores).

Dichos descriptores se agrupan en conjuntos semánticos según diferentes criterios y entre ellos guardan relaciones jerárquicas (términos genéricos o específicos) y asociativas.

Como vemos el papel del tesauro es fundamental. Es el elemento que sirve de puente entre el documentalista y el usuario, permitiendo que la búsqueda y recuperación de la información sea lo más eficaz posible, o lo que es lo mismo, exhaustiva (se recuperan todos los documentos del tema elegido) y precisa (sólo los de ese tema).

Sin el tesauro las búsquedas serían muy poco eficaces especialmente en el campo de las ciencias sociales y humanas donde lo habitual es que los términos sean muy polisémicos a diferencia de las ciencias exactas y naturales en que tales casos suelen ser las excepciones. A cualquiera se nos pueden ocurrir infinitud de ejemplos sobre este particular, por lo que no creo preciso abundar más en ello.

Un tesauro además de su carácter funcional, expresa el estado actual del desarrollo terminológico de una disciplina en un momento y lugar determinados según el marco socio-cultural imperante, e igualmente fija los modelos de clasificación e inter-relación entre los diferentes campos de cada materia.

Para concluir digamos que el tesauro es el lenguaje en el que se construyen una o unas bases de datos y mediante el cuál se accede a las mismas; es la infraestructura lógico-lingüística de un sistema documental.

Pasemos ahora a examinar la situación en el campo de las Ciencias de la Educación.

## LA SITUACIÓN DE LAS CIENCIAS DE LA EDUCACIÓN

Si nos ceñimos a la información aportada por la obra *Bases de Datos del Mundo* (1), las bases de datos centradas en el área educativa en general o en alguna sub-área son las siguientes:

- «Child Abuse and Neglect» (prevención y tratamiento de abusos y negligencias en la educación infantil).
- «ERIC» (Ciencias de la Educación en general).
- «Exceptional child education resources» (todos los aspectos de la educación de niños deficientes).
- «FRANCIS» (entre otras, Ciencias de la Educación en general).
- «NICSEM/NIMIS» (materiales para niños minusválidos).
- «Special Education Materials» (material impreso y audiovisual relativo a educación especial).

De este enumeración dos puntos me parecen especialmente significativos: en primer lugar que si exceptuamos la base FRANCIS todas son americanas, y en segundo lugar, de las seis citadas, cuatro se centran en temas relativos o cercanos a la educación especial.

En cualquier caso, y al margen de las conclusiones que cada uno quiera extraer, lo que parece a todas luces evidente es que tanto por su procedencia como por su temática, de una parte van a resultar insuficientes pues hay muchas cuestiones o áreas de trabajo que no van a aparecer, o bien parte de la información de que disponen no nos va a ser de utilidad por estar centrada en los modelos y pautas del ámbito cultural donde se produce (2).

Tomando en cuenta estas consideraciones y en función del volumen de publicaciones educativas que se editan en España o en castellano, parece cada vez más imprescindible poder disponer de bases de datos propios que pongan a disposición de todos la literatura científica que se produce, no sea que lleguemos a conocer antes y mejor las realizaciones americanas o del resto de Europa que las de nuestros más inmediatos vecinos.

Las realizaciones a este nivel son escasas: el Ministerio de Educación y Ciencia a través de la Secretaría General Técnica tiene disponibles las siguientes bases distribuidas por el centro de proceso de datos y accesibles mediante la RUE:

Se trata de las bases LEDA sobre legislación educativa, BIDE sobre cursos y becas de todo el mundo para estudiantes iberoamericanos, ISOC sobre Ciencias Sociales y Humanidades en general y REDINET actualmente en fase de implantación y a la que luego me referiré.

El área de disciplinas cubiertas es bastante restringido, sobre todo por el número de referencias almacenadas, pero más preocupante es la situación referida a los tesauros. Los existentes de los que tenemos noticia son:

- de tipo general; EUDISED, Ibe Education Thesaurus, Thesaurus Pädagogik, ERIC, Thesaurus of British Educational Terms, FRANCIS y Tesauro Colombiano de la educación.
- específicos; Tesauro de formación profesional, Evaluation Thesaurus, Thesaurus: enfance inadaptée, Special Education Theaurus, Terminology of adult education, Thesaurus education permanente.

De todo este conjunto los elaborados originalmente en lengua castellana son el colombiano y el de formación profesional de Uruguay.

Además de estos, los elaborados por organismos internacionales suelen ser tesauros multilingües, siendo normalmente el castellano uno de los idiomas de los mismos. Tal es el caso del EUDISED elaborado por el Consejo de Europa o el IBE por la UNESCO.

Pero en cualquier caso hay que tener presente que se trata usualmente de traducciones y no de tesauros específicamente adaptados a la realidad educativa y cultural de cada país. Son valiosos instrumentos para asegurar la comunicación internacional, pero en su propia concepción estriba también su principal defecto. A este respecto son clarificadoras las valoraciones realizadas sobre el de

la UNESCO por diferentes usuarios de los países latinoamericanos con sistemas educativos y terminologías muy variadas. (3).

Para suplir esta carencia son varias las iniciativas tomadas al respecto pero, que sepamos, hasta la fecha ninguna se ha plasmado en realizaciones concretas.

Como antes apuntamos el CSIC a través del ISOC tiene definido un plan de trabajo para elaborar tesauros castellanos de las diversas disciplinas de su competencia. Actualmente están elaborando uno de Urbanismo, y según nos han manifestado preven comenzar con el de educación no antes de dos años.

Junto a ello, autores como Rodríguez Delgado (4), se han manifestado en repetidas ocasiones partidarios de establecer planes conjuntos para la elaboración coordinada de tesauros en portugués y castellano.

Ya en la primera reunión de directores de centros de documentación educativa del área andina celebrada en Lima en 1978, se presentó el proyecto del «Tesauro educacional CENDIE» (Centro Nacional de Documentación e Información Educacional) con el patrocinio de la Organización de Estados Americanos.

Se realizó un estudio comparativo de los tesauros de la UNESCO, ERIC y EUDISED y se elaboraron una serie de descriptores a título experimental. Pero que sepamos, el proyecto no ha pasado de ahí.

Por último, en la sección de Pedagogía de la Facultat de Filosofía y Ciencias de la Educación de San Sebastián hemos comenzado un proyecto piloto centrado en la elaboración de un micro-tesauro referido al área de la investigación educativa, técnicas y metodología. Micro-tesauro que sería perfectamente integrable en un tesauro general de las Ciencias de la Educación en cuya elaboración participa el ICE de la Universidad del País Vasco y el centro de documentación del ITE de Alcalá de Henares, habiendo entablado conversaciones con el ISOC de cara a coordinar esfuerzos.

## **EL ÁREA DE LA INVESTIGACIÓN EDUCATIVA. TÉCNICAS Y METODOLOGÍA.**

Las razones de elegir este campo para la elaboración del primer micro-tesauro son varias; entre ellas –todo hay que decirlo– porque soy profesor del departamento de Pedagogía Experimental y es este el campo de las ciencias de la educación que menos desconozco. Pero al margen de ello, es preciso hacer notar que los que trabajamos en este campo, a la hora de realizar una búsqueda documental normalmente la realizamos porque nos interesa disponer de la información referente a lo trabajado en un tema determinado; pero no es menos frecuente que en lugar del contenido temático nos interese más disponer de las referencias de investigaciones que han desarrollado tal o cual método o diseño, éste o aquél procedimiento estadístico que nos interesa ver como opera independientemente de que sea una investigación sobre el aprendizaje lector o sobre la motivación de los alumnos.

Y para recuperar la información desde este punto de vista, los tesauros ac-

tualmente disponibles aportan un conjunto de descriptores muy restringido y poco adecuado.

Centrándonos en el EUDISED y en el de la UNESCO por ser éstos los más usuales en castellano, un cuadro comparativo sería más o menos el siguiente:

En el EUDISED de los 42 «conjuntos semánticos» que en que se halla dividido, el 16 corresponde a «Investigación». En total contiene 68 descriptores agrupados en cinco subconjuntos:

- *Métodos de investigación*, 21 descriptores como: «análisis comparativo», «método histórico», «análisis de items»...
- *Técnicas de investigación*, 19, tales como: «medición», «experimentación», «observación», «trabajo de campo»...
- *Tipos de investigación*, 16 como: «investigación aplicada», «investigación comparativa», «inv. empírica»...
- *Organización de la investigación*, 6 descriptores, entre ellos: «centro de investigación», «proyecto piloto»...
- *Método estadístico*, 6, a saber: «correlación», «covarianza», «factorial», «multivariado», «regresión» y «varianza».

Es de notar que de todos los métodos de investigación es el único que forma un grupo aparte.

El de la OIE-UNESCO considera el apartado investigación dentro del campo segundo con un total de 67 descriptores. Igualmente se subdivide en cinco:

- *Campos de investigación*, 10: «investigación curricular», «investigación del comportamiento», «pedagogía experimental»
- *Problemas*, 9: «prevención de la delincuencia», «problemas curriculares», «problema familiar».
- *Tipos de investigación*, 23: «análisis comparativo», «estudio de casos», «estudio estadístico»...
- *Metodología de investigación*, 15: «base de datos», «grupo de control», «validez»...
- *Datos*, 10: «análisis de datos», «análisis estadístico», «procesamiento de datos»...

Como puede verse si nos centramos en lo referido a la estadística, usando el tesoro EUDISED, sólo podemos distinguir entre cinco tipos de análisis, y en el de la UNESCO nos tendremos que contentar con el único término de «análisis estadístico» pues éste ni siquiera se subdivide.

En honor a la verdad hay que decir que no todos los tesoros son tan limitados como éstos, así la información aportada por ERIC es más amplia, pero repito que me he centrado en aquellos de los que existe versión castellana.

Desde el punto de vista del usuario esto supone que las búsquedas que se realicen van a ser de muy baja precisión, y que para obtener la información de-

seada vamos a vernos en la necesidad de extraer casi todas las referencias existentes agrupadas bajo dichos epígrafes. Cualquier resultado va a llevar implícito mucho «ruido», pues la mayoría de las referencias no van a ser pertinentes, con lo que ello lleva aparejado de costo adicional e ineficacia del sistema.

Evidentemente, la solución a esto es elaborar un tesoro, o mejor dicho, un micro-tesauro específicamente diseñado para cubrir y fijar la terminología de este ámbito con el nivel de precisión y exhaustividad necesario para solventar las necesidades planteadas.

Como vemos, centrados en el castellano y en las Ciencias de la Educación, de una parte las bases de datos existentes son pocas y cubren un espectro temático limitado, y de otra los tesauros existentes son traducciones de los multilingües resultando en ocasiones inadecuados.

Como un intento de contribuir a ir remediando esta situación, se inscribe el proyecto REDINET del que seguidamente paso a describir sus características más importantes.

## EL PROYECTO REDINET. CARACTERÍSTICAS GENERALES.

Teniendo en cuenta las cuestiones antes formuladas y considerando además el continuo incremento del número de investigaciones sobre educación que en España se realizan, el CIDE (organismo impulsor del proyecto) consideró que era necesario instrumentar un sistema de información que asegurase la mutua comunicación entre todos los investigadores. Necesidad que se ve más justificada al considerar la diversidad de instituciones que producen o financian trabajos de este tipo (Facultades y Secciones de Pedagogía, Escuelas de formación del Profesorado, ICEs, Consejerías de Educación de las Comunidades Autónomas...).

La primera decisión fue la referida a estructurar la organización según el modelo de las Comunidades Autónomas, de tal manera que más que de una base de datos, hay que hablar de una red de bases de datos integrada por el CIDE y las 17 Consejerías de Educación o instituciones en que éstas hayan delegado.

Se considera que cada miembro de la Red es autónomo en su demarcación, pudiendo, si su infraestructura y medios lo permiten, formar a su vez su propio servicio de documentación.

Para asegurar el principio de reciprocidad e igualdad entre los miembros, cada uno se compromete a remitir al Centro de Proceso de Datos del Ministerio la información documental referida a las investigaciones producidas en su zona con el fin de que esté disponible para el resto.

Se podría hablar pues de un servicio de documentación descentralizado, o mejor semi-centralizado, en la medida en que los integrantes del sistema no se limitan a priori a ser únicamente terminales del ordenador central, aunque por obvias razones económicas es éste el que almacena la totalidad de la documentación generada por cada miembro poniéndola a disposición de toda la Red.

De esta manera cada miembro adecúa su modo de operar a los fondos de que dispone y a las necesidades y demandas que se le planteen, de forma que habrá centros a los que les compensará más operar como terminales del ordenador central en la mayoría de los casos, mientras que en otros el funcionamiento será normalmente autónomo.

Como contrapartida esto exige que la infraestructura informática sea compatible y sobre todo disponer de formatos y lenguajes documentales comunes a todos para asegurar las intercomunicación, máxime si consideramos que si bien el ordenador central dispone de todas las referencias enviadas por cada miembro, las fuentes primarias se encuentran solamente allá donde se han producido, de tal manera que una búsqueda documental debe proporcionar además de la información habitual, la necesaria para que el usuario pueda localizar los documentos que le interesen.

Centrándonos ahora en lo referido a los formatos, la ficha para realizar el vaciado e «indización» documental consta de dos partes:

La primera contiene los datos de información general para identificar y localizar el documento original (título, autor, datos editoriales si está publicado o datos del centro donde es localizable en caso contrario etc).

Los campos de la segunda parte se basan en lo ya comentado. Tan útil es la información temática de un trabajo como la referida a las técnicas y metodología del mismo. Así se incluyen campos como los objetivos de la investigación, la muestra o el objeto del trabajo, los instrumentos de obtención de información, las técnicas de análisis etc, concluyendo con los descriptores temáticos y metodológicos.

Esta ficha es el primer elemento de la necesaria normalización de la Red. Por lo que se refiere al segundo, la comisión asesora optó por utilizar en una primera fase el tesoro EUDISED pues pensamos que su terminología y su sistema de clasificación de las familias semánticas resultaban más adecuados.

En cualquier caso ya hemos comentado que presenta problemas, y entre ellos el referido a las cuestiones metodológicas. Para remediar en lo posible esta cuestión actualmente se dispone de listas controladas de términos que provisionalmente pueden ser útiles hasta que no se disponga del o de los tesauros necesarios.

Por último señalar que el logical que soporta el sistema documental es el paquete UNIDAS desarrollado por la casa UNIVAC. Este permite recuperar documentos a través de tres áreas:

- a) aspectos significativos del documento (lo que hemos llamado información de tipo general: autor, tipo de investigación, año).
- b) descriptores (en nuestro caso de ambos tipos).
- c) búsqueda en texto libre (todos los campos destinados a la descripción pormenorizada de las características metodológicas).

En cualquiera de las tres formas de recuperación de la información o en sus posibles combinaciones se acota el conjunto documental mediante los operadores booleanos habituales.

## CONCLUSIONES FINALES

Referido a la documentación científica automatizada en el campo de las Ciencias de la Educación se observa el bajo número de bases de datos existentes con un campo temático limitado y la inexistencia de tesauros o micro-tesauros elaborados originalmente en lengua castellana y adecuados al marco socio-cultural español.

Esto conlleva una fuerte dependencia externa y un bajo nivel de comunicación entre los miembros de la comunidad científica. (Como nota complementaria digamos que nos ha sido imposible disponer de los tesauros colombiano o uruguayo y en general tener conocimiento preciso de las realizaciones y proyectos planteados en Latinoamérica).

La primera línea de acción para solventar estas carencias pasa por aumentar el nivel de comunicación entre los investigadores con el fin de lograr un acuerdo y poder disponer de un lenguaje documental común por mínimo que éste sea.

El estado actual de la tecnología informática y en general de la investigación sobre tratamiento automático de la información implica el uso de lenguajes controlados y normalizados, y mientras esto no cambie lo anterior es un requisito previo para cualquier acción ulterior.

Logrado esto sería posible generar pequeñas bases de datos locales o sectoriales fácilmente interconectables entre si a un costo relativamente bajo, pues el desarrollo de la micro-informática lo posibilita.

## NOTAS

1. FUINCA, (1981) *Bases de Datos del Mundo*. Alhambra. Madrid.
2. VILLAGRÁ, A., y ROMÁN, A., (1981). «Diferente utilidad de las bases de datos americanas y europeas en las ciencias sociales» *Rev. Esp. Doc. Cient.* Vol. 4, Número 2, pp. 112-129.
3. Cfr. Los siguientes artículos publicados en la *Revista de la UNESCO de Ciencias de la Información, Bibliotecología y Archivística*:

BRAHM, L., y GUTIÉRREZ, G., (1983): «REDUC: una red de documentación en educación al servicio del desarrollo en América Latina». Vol. V, Número 2, Abril-Junio. pp. 93-97.

- FURIADO, A., (1983) «La Red Internacional de Información Educativa (INED) y la participación de America Latina y el Caribe». *Ibid.* pp. 86-89.
4. RODRÍGUEZ DELGADO, R., (1979): «Plan para la elaboración de Tesauros coordinados en portugués y español». *Rev. Esp. Doc. Cient.* Vo. 2, n.º 4, pp. 337-355
- RODRÍGUEZ DELGADO, R., (1980).: «La integración de las lenguajes documentarios, fin de Babel». *Rev. Esp. Doc. Cient.*, Vol. 4, n.º 3, pp. 333-341.



## librería herder

### SERVICIO INTERNACIONAL

les ofrece su sección de:

- **Psicología**
- **Pedagogía**
- **Psicoanálisis**
- **Tests**

Con unos 5.000 títulos en stock permanente

Solicite el envío gratuito del boletín  
bibiográfico de estos temas o, mejor,  
visite nuestra exposición.

**LIBRERIA HERDER**  
Balmes, 26 - Tel. 317 05 78  
08007 BARCELONA

# Meta-análisis sobre la predicción del rendimiento escolar

por  
M.ª Paz Loscos

## OBJETIVOS DEL TRABAJO

La importancia que está tomando en el campo de la investigación pedagógica el tema de la predicción del rendimiento escolar y el considerable aumento producido en los últimos años de su literatura, provocan el hecho de adoptar la técnica del meta-análisis en un intento de clarificar la situación actual del tema, pues si bien son muy numerosas las investigaciones realizadas por psicólogos, pedagogos y maestros para un mismo tema de trabajo, todas ellas difieren en algún aspecto, como pueden ser las características de la muestra, el número de variables predictoras que intervienen, la combinación de las mismas, los métodos de predicción utilizados e incluso, valga la redundancia, el criterio adoptado por cada autor para valorar la variable «criterio».

Todo ello hace que sea muy difícil hacer una agrupación de estudios de similares características con el fin de poder emitir un juicio generalizable derivado de unas conclusiones determinadas.

Es por esto, que se adoptó la decisión de aplicar las técnicas del meta-análisis a estos estudios para que, al analizar sus resultados estadísticamente a través de sus variables y características, pudiera llegarse al conocimiento de la situación real y actual del problema, considerando como fin de trabajo en estas líneas, pero como medio que, a través del cual, permita a los profesionales de la educación fijar o alterar los objetivos y estrategias desarrolladas en el sistema escolar.

## DESARROLLO DEL PROBLEMA

### 1. Definición de conceptos

- *Meta-Análisis:* Es una técnica de integración cuantitativa que relaciona esta-

dísticamente las características de diseño para un determinado tipo de estudios. Según Glass, G.V. (1978) es «el análisis del análisis».

- *Predicción:* En el campo de las Ciencias Humanas, supone el efecto de anticipar situaciones o consecuencias en las que podría basarse el comportamiento humano.
- *Rendimiento:* Determina el nivel alcanzado por el grado de dificultad en la resolución de ciertas pruebas. Su valoración cuantitativa proporciona información para realizar diagnósticos longitudinales.

## 2. Métodos predictivos

- *Clinico:* Se basa en la observación y opinión subjetiva del orientador (Lindzey, G., 1965; Goldman, L., 1972).
- *Estadístico:* Predice una determinada conducta a través de unas puntuaciones obtenidas mediante la medición de ciertas variables independientes o predictoras (Meehl, P., 1956).
- *Mixto:* Es una integración de los dos métodos anteriores. (Holt, R.R., 1958; Castaño López-Mesas, C., 1974; F. Ballesteros, R., 1980).

## 3. Ecuaciones de Predicción

Si se considera como la variable criterio el rendimiento académico, actuarán como predictores una serie de variables tanto de tipo aptitudinal como de personalidad y pedagógicas; teniendo que determinar la validez de las asociaciones establecidas entre variables mediante técnicas de correlación múltiple, como base de las ecuaciones de predicción, para poder establecer una inferencia causal a través de la cual se pueda predecir qué nivel de rendimiento puede conseguir un sujeto a través de las puntuaciones obtenidas en las variables predictoras, expresado por el coeficiente de correlación múltiple R.

## RESEÑA HISTÓRICO-BIBLIOGRÁFICA

### 1. Antecedentes teóricos

El tema de la predicción del rendimiento escolar expresado mediante el coeficiente de correlación múltiple R es relativamente moderno dentro del campo de las Ciencias de la Educación.

Se produce un acercamiento a este tipo de predicción en autores como Gali, A., (1928) con sus «coeficientes de eficiencia», la polaca María Kaczynska (1935); y ya más adelante Fernández Huerta, J., (1947) y García Hoz, V., (1952),

en los que se percibe una predicción teórica que tendría que ser reforzada estadísticamente.

## 2. Estudios de predicción estadística

1950... Aparecen en España los primeros estudios de predicción escolar con Francisco Secadas en 1952, elaborados estadísticamente mediante ecuaciones de regresión múltiple, con las variables predictoras de inteligencia general y personalidad. ( $R = 0.43$ ).

1960... García Yagüe, J., López Aberástury, N., Velasco, I., Urcullu, B., y Caballero, A., añaden a los factores predictores de Secadas otros de tipo pedagógico, (0.710-0.800).

1970... Prosiguen en esta línea de investigación Pacheco del Olmo, B., Palomino López-Menchero, H., y Crespo, J., (R hasta 0.808). Rivas, F., realiza un extenso estudio que abarca toda la E.G.B. ( $R$  promedio = 0.7879). En Zaragoza, el equipo de Cid Palacios, R., Bernad, A., y Escudero, T., realizan estudios de predicción para los alumnos de Acceso a la Universidad, con un seguimiento de 5 años ( $R = 0.53$ ).

1980... Chorro Gascó, J.L., Fernández Dols, J.M., Salvador Rodríguez, M.; todos ellos psicólogos, ponen de manifiesto el interés que también despierta el tema en el campo psicológico. Se encuentran interesantes estudios con factores predictores de tipo lingüístico de Molina García ( $R = 0.49-0.60$ ) y Rodríguez Diéguez, J.L., ( $R = 0.696-0.716$ ). Resulta notable la introducción de nuevos predictores de hábitos de estudio, autoconcepto, popularidad, integración social, sinceridad, etc.

## CARACTERÍSTICAS DE LOS ESTUDIOS

### 1. Determinación de variables

Con el fin de realizar una síntesis de datos que pudieran ser tratados estadísticamente, se han ido tomando los estudios agrupados de la siguiente forma, según las distintas variables a considerar:

#### - *Variables dependientes o «criterios»*

- Rendimiento escolar general.
- Rendimiento escolar en Ciencias.
- Rendimiento escolar en Letras.

- *Variables independientes o «predictores»*

- Inteligencia general, factor «g» (a).
- Aptitudes diferenciadas, (b).
- Factores de personalidad, (c).
- Factores pedagógicos, (d).
- Rendimiento anterior, (e).
- Popularidad e integración social, (f).
- Otras (autoconcepto, sinceridad, hábitos y técnicas de estudio, etc.).
- Combinación de factores
  - a-b, a-c, y a-d.
  - a-b-c, a-b-d, y a-b-e.

- *Variables de identificación*

- Muestra: ,Edad.  
,Nivel escolar.  
,Origen.
- Artículo: ,Fuente de publicación.  
,Fecha de publicación.  
,Lugar de publicación.  
,Autores.

- *Otras variables*

- Tests empleados: ,Inteligencia general.  
,Aptitudes diferenciadas.  
,Pruebas pedagógicas.

## 2. Síntesis numérica

Para la realización de este trabajo han sido seleccionados 50 estudios de una pre-selección anterior de 285; pues sólo el 17,54% eran válidos para el metaanálisis pro su condición indispensable de expresar sus resultados en coeficientes de correlación múltiple R.

Se han obtenido de esos 50 estudios, 271 datos en términos de coeficientes de correlación múltiple R, lo cual da una media de 5,42 datos por estudio. Estos 271 datos alcanzan valores entre 0.14 y 0.914, con una amplitud de variable de 0.764.

Todos estos datos han sido extraídos de estudios publicados en revistas de pedagogía y psicología, tesis y tesinas, libros, actas de congresos y publicaciones del Ministerio de Educación y Ciencias (I.C.E. e I.N.C.I.E.).

La amplitud temporal de búsqueda de datos abarca 40 años, si bien la real se redujo a 32:

E. localizados ... 1945-1985 --- 40 años  
E. seleccionados ... 1952-1984 --- 32 años

De los 28 autores que componen estos estudios, alcanzan la mayor frecuencia Secadas, F., con 8 artículos; Crespo, J., con 5; y García Yagüe, J., 5.

El 95% de los estudios seleccionados, lo han sido en bibliotecas de especialización pedagógica, como el Instituto de San José de Calasans (C.S.I.C.), el Instituto de Psicología Aplicada y Orientación, Instituto de Educación Especial, Facultades de Psicología y Pedagogía, I.N.C.I.E., M.<sup>o</sup> de Educación y Ciencia, etc.

La falta de datos sobre las muestras de 15 estudios ha sesgado notablemente algunos resultados y ha hecho imposible realizar otras clasificaciones importantes de acuerdo con el sexo, nivel socio-cultura, etc.

## ANÁLISIS ESTADÍSTICO

### 1. Aplicación de los estadísticos

Se presenta primeramente el resultado general del metá-análisis obtenido en la resolución de los estadísticos de todos los coeficientes de correlación múltiple R, mediante ecuaciones de regresión múltiple entre las variables predictoras relacionadas y la variable criterio del rendimiento escolar; siempre que se considere la predicción de éste en su más amplio sentido, como una unificación del concepto predictivo, independientemente de la distribución de sus variables.

Con estos coeficientes, se ha procedido a hallar el valor de su media, mediana, desviación típica, error típico y amplitud de su variable, así como la medida del porcentaje de varianza.

A continuación, la resolución de estadísticos se aplica a los resultados en R que figuran en los estudios seleccionados para la variable que interesa relacionar, según la clasificación efectuada en el punto 4.1. (determinación de variables).

A tal efecto, se presentan unas tablas en las que figura, por columnas, junto a la variable a considerar:

- N (E): número de estudios en los que aparecen valores de coeficientes de correlación múltiple R para esa variable.
- N (R): número de datos en R que contienen esos estudios.
- X (R): valor medio de todos los datos R, de la variable.
- $\sigma$  (R): desviación típica todos los datos R, de la variable.
- $E\sigma$  (R): error típico de todos los datos R, de la variable.
- X ( $R^2$ ): valor medio de los datos R elevados al cuadrado.
- X (%): valor medio de porcentaje de varianza explicada para la variable.

## 2. Tablas de resultados

### - Variables dependientes

#### T.1.: Rendimiento escolar general

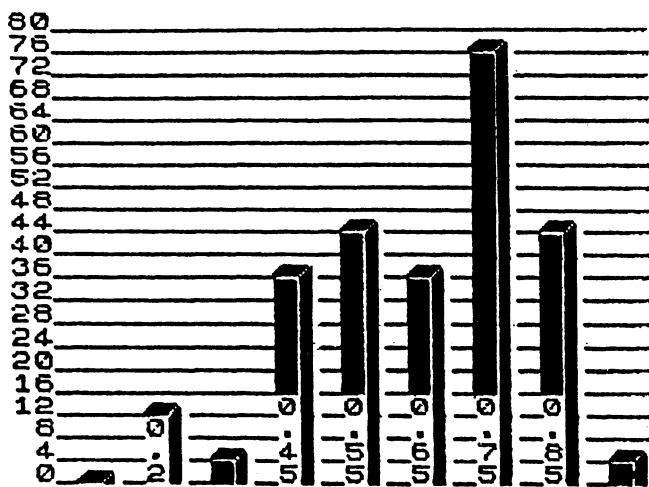
##### *R. Numéricos:*

N.º de estudios	:	NE	= 50
N.º de datos	:	NR	= 271
Media de los datos	:	$\bar{X}R$	= 0.639
Mediana de los datos	:	Mdn.	= 0.680
Desv. típica de datos	:	$\sigma$	= 0.172
Error típico de datos	:	E	= 0.010
Dato menor	:	R	= 0.14
Dato mayor	:	R	= 0.914
Amplitud variable	:	Av.	= 0.764
Media de cuadrado de			
datos	:	$\bar{X}^2$	= 0.437
Media de varianza	:	$\bar{X}\%$	= 43.7

##### *R. Gráficos:*

Asimetría negativa: As.= -0,715

#### ESTADÍSTICA GENERAL



## T.2.: Rendimiento escolar por áreas

Áreas	E		R	
	N	N		
Ciencias	12	48		
Letras	7	16		
Total	19	64		

-Variables independientes

## T.3.: Factores predictores aislados

Factores	E		R		$R^2$		%
	N	N	$\bar{X}$	$\sigma$	$E\sigma$	$\bar{X}$	$\bar{X}$
a) Inteligencia general, f. «g»	14	47	0.633	0.137	0.020	0.4202	42.02
b) Aptitudes diferenciales	12	23	0.540	0.178	0.037	0.3238	32.38
c) Factores de personalidad	5	14	0.499	0.068	0.018	0.2539	25.39
d) Factores pedagógicos	7	14	0.544	0.212	0.057	0.3410	34.10
e) Rendimiento anterior	8	10	0.617	0.163	0.052	0.4073	40.73
f) Popularidad e integración social	3	3	0.621	0.061	0.035	0.3890	38.90
Total	49	111					

## T.4.: Combinación de 2 factores

Factores	E		R		$R^2$		%
	N	N	$\bar{X}$	$\sigma$	$E\sigma$	$\bar{X}$	$\bar{X}$
a)-b) Int. gral. «g» Ap. difer.	6	8	0.716	0.104	0.037	0.5233	52.33
a)-c) Int. gral. «g» Fc. person.	4	4	0.502	0.117	0.059	0.2663	26.63
a)-d) Int. gral. «g» Fc. pedag.	9	22	0.759	0.067	0.014	0.5801	58.01
Total	19	34					

T.5.: Combinación de 3 factores

Factores	E		R		$R^2$		% $\bar{X}$	
	N	N	$\bar{X}$	$\sigma$	$E\sigma$	$\bar{X}$	$\bar{X}$	$\bar{X}$
a)-b)-c) Ap. ment. Fc. pers.	6	6	0.542	0.119	0.048	0.3084	30.84	
a)-b)-d) Ap. ment. Fc.pedag.	13	26	0.760	0.063	0.012	0.5811	58.11	
a)-b)-e) Ap. ment. Rend. ant.	6	30	0.722	0.080	0.014	0.5278	52.78	
Total	25	62						

– *Variable de identificación*

• *Muestra*

T.6.: Edad

T.7.: Nivel escolar

Años	E		R	
	N	N	N	N
6 – 10	14	45		
11 – 15	28	120		
16 – 20	9	66		
21 – 25	3	21		
Total	54	252		

Nivel	E		R	
	N	N	N	N
E.G.B. (1. <sup>a</sup> Etapa)	13	44		
E.G.B. (2. <sup>a</sup> Etapa)	25	94		
B.U.P.	6	32		
C.O.U.	3	18		
Universidad	5	42		
Total	52	230		

T.8.: Origen

Origen	E		R	
	N	N	N	N
Madrid	26	102		
Provincias	17	121		
Total	43	223		

●Artículos

T.9.: Fuente de publicación

Publicaciones	E	R
	— N	— N
Actas de Congresos	6	23
Libros y manuales	3	21
Tesis doc. y tesinas	3	6
Publicaciones M.º E. y C. (I.C.E. e I.N.C.I.E.)	6	60
Revistas de Pedagogía	20	116
Revistas de Psicología	12	45
Total	50	271

T.10.: Fecha de publicación

Años	E	R
	— N	— N
1945 – 1954	3	7
1955 – 1964	7	37
1965 – 1974	12	47
1975 – 1984	28	180
Total	50	271

T.11.: Lugar de publicación

Lugar	E	R
	— N	— N
Madrid	36	195
Provincias	14	76
Total	50	271

T.12: Autores

Autor	E	R
	— N	— N
Escudero	4	36
García Yagüe	4	26
Palomino	4	24
Secadas	8	35
Total	20	121

- *Otras variables*

• *Tests empleados*

T.13.: Inteligencia general, factor «g»

Test	E		R	
	N		N	
D-48	6		28	
OTIS	7		20	
BALLARD	10		25	
A.M.P.E.	7		20	
Aptitudes generales (García Yagüe)	9		25	
Total	39		118	

T.14: Aptitudes diferenciales

Test	E		R	
	N		N	
D.A.T.	5		24	
A.M.D.I.	6		8	
P.M.A.	5		10	
Total	16		42	

T.15.: Pruebas pedagógicas

Pruebas	E		R	
	N		N	
Publicadas	15		72	
Elaboración personal	14		94	
Total	29		166	

## INTERPRETACIÓN DE LOS RESULTADOS

### 1. Consideraciones previas

En la realización del meta-análisis sobre la predicción del rendimiento escolar se ha trabajado con 50 estudios que, mediante ecuaciones de regresión múltiple, han proporcionado 271 coeficientes de correlación múltiple ( $R$ ), los cuales deben ser considerados como estimaciones del aumento o disminución de la variable dependiente (en este caso el rendimiento escolar), originados por el aumento o disminución de una unidad en la variable independiente (inteligencia, personalidad, rendimiento anterior, autoconcepto, etc.).

Por tanto, estos coeficientes son predicciones que, con un cierto nivel de error, informan sobre el grado de aportación de una variable a la predicción del valor de otra y dependiendo siempre de las variables introducidas.

Esta posibilidad de predicción no debe tomarse nunca como una causalidad, sino como una determinación por la variable independiente o por otras variables no tenidas en cuenta en el estudio pero relacionadas con la variable en cuestión. Es decir, que si en un estudio aparece una variable con un gran peso en la ecuación de regresión, hará posible que mediante esa variable pueda predecirse para la población que la muestra representa, la variación media de la variable dependiente, en el caso de que las demás variables permanezcan constantes.

Como es algo difícil que pueda darse esta condición en la realidad, de igual distribución de la variable en las muestras, no pueden hacerse comparaciones de sus resultados predictivos, ni hacerlos extensivos para otra población.

### 2. Comentarios a las tablas de resultados

#### – *Variables dependientes*

##### • *Rendimiento escolar general*

El valor medio de 0.639 de todos los coeficientes expresados en  $R$  ( $\frac{X}{N}$ ) que se incluyen en los 50 estudios es bastante significativo, si además se tiene en cuenta la gran dispersión que presenta los 271 datos en la gama de variable 0.764, que abarca valores desde 0.14 a 0.914; tal y como puede apreciarse en la representación gráfica de su distribución, en la que puede distinguirse igualmente que dicha distribución de los datos figura sesgada negativamente (-0.715).

También se ha hallado la media de la varianza total de cada uno de los 271 datos elevados al cuadrado ( $\frac{X^2}{N}$ ) obteniéndose un porcentaje medio del 43.7% ( $0.437 \times 100$ ). No se ha elevado al cuadrado el dato de media obtenido de 0.639 porque hubiera dado una media de  $R^2$  de 0.408, con una varianza media de 40,8%, inferior en un 2.9%. (Tabla 1).

- *Rendimiento escolar por áreas*

Si se subdivide la variable criterio de rendimiento escolar en las áreas de Ciencias y Letras, solo 19 de los 50 estudios aportan 64 datos para realizar esta clasificación (23.62% del total de 271), siendo además el triple para Ciencias que para Letras (NR = 48 y 16, respectivamente).

Como en estos 19 estudios no se mantenían las mismas variables independientes, no se han efectuado cálculos estadísticos de media, mediana, desviación típica, etc., porque hubiera quedado muy disminuido el número de artículos con una constante de factores predictivos. (T.2.)

- *Variables independientes*

Se han tomado los valores en R de cada factor que era considerado como variable predictor o independiente en cada estudio en el que aparecía como valor único para esa variable.

- *Inteligencia*: se ha subdividido en inteligencia general, factor «g» (que alcanza el mayor nivel predictivo medio: 0.633) y aptitudes diferenciales (0.540).
- *Personalidad*: este factor explica el menor porcentaje medio de varianza: 25.39%.
- *Factores pedagógicos*: su nivel predictivo medio supera al de las aptitudes diferenciales minimamente (0.544).
- *Rendimiento anterior*: alcanza también un alto valor predictivo (0.617).
- *Popularidad e integración social*: a pesar de su escasa frecuencia resulta notable su porcentaje medio de varianza, un 38.90% (T. 3).

- *Combinación de variables independientes*

Como no todos los autores combinan los mismos factores predictivos, se han encontrado pocos estudios que ofrecieran las mismas constantes de conjuntos factoriales, por lo que sólo ha sido posible realizar el cálculo estadístico con la integración de factores que aparecen en las tablas.

- *2 factores*

Se combina el factor de inteligencia general, factor «g», con los de aptitudes diferenciales, de personalidad o pedagógicos. Resalta el alto valor medio predictivo alcanzado por la unión de la inteligencia general, f. «g» y factores pedagógicos (0.759), seguida de la unión de inteligencia general y aptitudes diferenciales (0.716). Ambas modalidades rebasan el 50% medio de varianza (58.01% y 52.3%, respectivamente) y suponen un doble porcentaje frente al alcanzado por la inteligencia general y los factores de personalidad (0.52 26.63%). (T. 4.).

- *3 factores*

Permanecen constantes las variables de inteligencia general, factor «g» y aptitudes diferenciales (que para una mayor simplificación se considera ambas bajo el título de aptitudes mentales) y se añade el factor de rendimiento anterior, que alcanza la mayor frecuencia unido a las aptitudes mentales. Comparando brevemente esta tabla combinatoria de 3 factores con la anterior de 2, se observa una cierta correspondencia de los resultados, según los factores a combinar, ya que la media superior de los valores del coeficiente R se alcanza cuando se unen las aptitudes mentales con las calificaciones obtenidas en cursos anteriores (0.722) y el valor medio inferior corresponde a las aptitudes mentales y la personalidad (0.542). (T.5.).

- *Variables de identificación*

Por la limitación del número de estudios encontrados para el meta-análisis y su característica apuntada anteriormente de dispersión en cuanto a la elección del número de factores predictivos y la combinación de éstos, no ha sido posible hallar para cada subvariable los mismos factores de predicción en número suficiente como para constituir una muestra representativa con la que pudieran formularse ciertas hipótesis de predicción. Se presentan, por tanto, solo las tablas con las características de las muestras y de los artículos, con los datos del número de estudios y de los coeficientes de correlación R. que contienen. (Tablas 6, a 12.).

- *Otras variables*

Respecto a los tests empleados en cada uno de los estudios tampoco ha podido efectuarse el meta-análisis, por la misma razón expuesta para las variables de identificación. Sin embargo, se ha creido interesante realizar una clasificación de los tipos de tests que se utilizaban para medir cada factor predictivo, especificando número de estudios y de coeficientes R que arrojaban mediante su aplicación. (T. 13 a 15).

## SÍNTESIS

Desde que en el año 1952 Secadas, F., realiza el primer estudio en España sobre predicción del rendimiento escolar, tomando como variables predictoras factores de inteligencia y personalidad en las ecuaciones de regresión múltiple, casi todos los autores que han seguido después investigando sobre este tema (incluido el mismo Secadas) han introducido otros factores como los pedagógicos,

el rendimiento anterior, autoconcepto, hábitos de estudio, técnicas de base, popularidad e integración social, ambiente socio-económico, etc. Estos factores han sido tomados unas veces independientemente y otras realizando combinaciones de 2, 3, 4 ó todos ellos juntos; quedando patente el aumento del valor del coeficiente de correlación R, a medida que se incrementa el número de variables predictoras en las ecuaciones de regresión múltiple.

Es por esto que se creyó sumamente interesante realizar una recopilación de todos los estudios publicados en España respecto a este tipo de investigación para aplicarles las técnicas del meta-análisis. Para ello fueron seleccionados 50 estudios que expresaban sus predicciones escolares en términos de coeficientes de correlación múltiple R, como condición indispensable para poder realizar una integración de resultados.

Se efectuaron primeramente los cálculos estadísticos de todos los resultados obtenidos sobre predicción escolar que dieron para un total de 271 coeficientes R, una media de 0.639. Después se pasó a realizar una integración de los resultados por variables, alcanzando el mayor valor predictivo medio la inteligencia general, factor «g» (0.633) y en el resultado de la combinación de variables predictivas, la unión de los factores pedagógicos con la inteligencia general (0.759).

No obstante, no deben tomarse los resultados obtenidos ni como una razón de causalidad de sus variables, ni como una generalización a una población distinta a la que representa la muestra.

#### REFERENCIAS BIBLIOGRÁFICAS

- CASTAÑO LÓPEZ-MESAS, C., (1974): «Problemas de predicción de los tests mentales en referencia a la orientación escolar y profesional». *Revista de Psicología General y Aplicada*, v. XXIX, 126, pp. 31-41.
- CID PALACIOS, R., y colbs. (1976): *Rasgos de madurez y éxito en las pruebas de acceso a la Universidad*. I.C.E. Zaragoza.
- CRESPO VASCO, J., (1971): «El problema de la estabilidad de los tests». *Educadores*, v. XIII, 61, pp. 401-417.
- CHORRO GASCO, J.L., (1981): «Personalidad y rendimiento académico en Estadística». *Psicológica*, v. II, 2, pp. 155-166.
- ESCUDERO ESCORZA, T., y colbs. (1981): *Racionalidad del proceso de selección de universitarios: planteamiento de un modelo*. I.C.E. Zaragoza.
- FERNÁNDEZ BALLESTEROS, R., (1980): *Psicodiagnóstico. Concepto y metodología*. Cincel-Kapelusz. Madrid.
- FERNÁNDEZ DOLS, J.M. y SERRANO MARTÍNEZ, G., (1980): «Socialización y rendimiento académico en la mujer». *Psicológica*, v. I, 2, pp. 215-223.
- FERNÁNDEZ HUERTA, J. (1947): «Investigación pedagógica. Procedimientos de evaluación». *Revista Española de Pedagogía*, v. V, 18, pp. 275-287.
- GALI, A., (1928): *La medida objetiva del trabajo escolar*. Aguilar. Madrid.

- GARCÍA HOZ, V., (1952): «Nociones y aspectos del rendimiento escolar». *Bordon*, v. IV, 25-26, pp. 3-8.
- GARCÍA YAGÜE, J., y LÓPEZ ABERASTURI, N., (1964): «El pronóstico para los estudios del bachillerato elemental al nivel de ingreso». *Revista de Psicología General y Aplicada*, v. XIX, 73, pp. 523-526.
- GOLDMAN, L., (1972): «Tests and counseling: The marriage that failed». *Measurement and Evaluation in Guidance*, 4, pp. 213-220.
- GLASS, G.V., (1978): «Integrating findings: The meta-analysis of research». *Review of Research in Education*, V, Itasca III. Peacock.
- HOLT, R.R., (1958): «Predicción clínica y estadística: una reformulación y nuevos datos». *Journal of abn. and Soc. Psych.*, 56, pp. 1-12.
- KACZYNSKA, M., (1935): *El rendimiento escolar y la inteligencia*. Espasa-Calpe. Madrid.
- LINZDZEY, G., (1965): «Vidente en contraposición a signo». *Journal of Experimental Research in Personality*, 1, pp. 17-26.
- MEEHL, P., (1954): *Clinical vs. Statistical Prediction*. University of Minnesota Press. Minneapolis.
- MOLINA GARCÍA, S., (1984): *Códigos lingüísticos y rendimiento escolar*. VIII Congreso Nacional de Pedagogía. Santiago de Compostela.
- PACHECO DEL OLMO, B., y CABALLERO, A., (1977): *Diagnóstico del rendimiento pedagógico en E.G.B.* Miñón. Madrid.
- PALOMINO LÓPEZ-MENCHERO, A., (1970): «La predicción del éxito en el bachillerato superior». *Educadores*, v. XII, 57, pp. 203-220.
- RIVAS MARTÍNEZ, F., (1977): «Orientación y predicción escolar» *Vida Escolar*, v. XIX, 191-192, pp. 67-72.
- RODRÍGUEZ DIEGUEZ, J.L., (1984): *Códigos lingüísticos y procesos de enseñanza*. VIII Congreso Nacional de Pedagogía. Santiago de Compostela.
- SALVADOR RODRÍGUEZ, M., (1981): «Análisis experimental de resultados académicos y formativos en la 2.<sup>a</sup> etapa de E.G.B.». *Psicológica*, v. II, 1, pp. 23-40.
- SECADAS MARCOS, F., (1964): «Discrimen de aptitud para estudios medios o de aprendizaje». *Revista Española de Pedagogía*, v. XXII, 86, pp. 91-115.
- SECADAS MARCOS F., (1952): «Factores de personalidad y rendimiento escolar». *Revista Española de Pedagogía*, v. X, 37, pp. 77-87.
- URCULLU, B., VELASCO, I., LÓPEZ ABERASTURI, N., y MARTÍN, B., (1964): «Primeros resultados de una batería psicopedagógica». *Revista de Psicología General y Aplicada*, v. XIX, 73, pp. 481-498.

## Meta-análisis correlacional sobre estudios de rendimiento escolar en España

por  
Juan Mateo

### INTRODUCCIÓN

La investigación educativa produce anualmente, a nivel internacional, cientos de estudios sobre un mismo tópico. Sin embargo nuestro estado del conocimiento permanece, en muchas ocasiones, estacionario, a pesar de los ingentes esfuerzos de la comunidad científica, para profundizar en la comprensión de dichos tópicos o al menos para resolver los problemas que de su ignorancia puedan derivarse.

Las causas de tanta esterilidad son diversas y no es el caso de entrar aquí en un análisis de las mismas, aunque si vamos a plantearnos desde una perspectiva metodológica una revisión de las técnicas cuantitativas agrupadas en torno al llamado meta-análisis, que al permitirnos integrar los logros obtenidos en diversos estudios sobre un mismo centro de interés, aportará un grano de arena en la superación del estaticismo a que aludíamos al principio.

Complementaremos la revisión teórica con una aplicación práctica a una serie de estudios correspondientes a los últimos años, centrados en la aportación de las variables aptitudinales en la explicación del rendimiento escolar.

Los conceptos básicos que subyacen al meta-análisis ya fueron empleados por Thorndike, R.L., (1933) y Ghiselli, G., (1949) y más recientemente por Fleishman, L., y Levin, K.L., (1973, 1975) pero podemos quizás considerar a Light, J., y Smith, M.L., (1971) los pioneros en examinar de forma sistemática, al problema de integrar los resultados de diferentes estudios cuantitativos en el campo de las Ciencias Sociales. En contrapartida diremos que la mayoría de sus procedimientos requieren tener acceso a los datos originales, hecho que ha dificultado enormemente su aplicación y divulgación.

Sin embargo, no es hasta muy reciente que no aparecen y se populariza el uso de distintas técnicas cuantitativas agrupadas de forma sistemática. Es Glass, G.V., (1976) el primero en avanzar un grupo de ellas y a él debemos también el

término «meta-análisis» entendido como «el análisis de los análisis» término que fue así acuñado para distinguirlo del llamado «análisis secundario». En un análisis secundario el investigador obtiene y re-analiza los datos originales sobre los que se basó un estudio anterior (Light, J., y Smith, M.L., 1971) mientras que el meta-análisis, en el sentido que le confiere Glass, *es la acumulación cuantitativa y análisis de estadísticos descriptivos recogidos a través de distintos estudios*, y no requieren tener acceso a los datos originales.

Desde su aparición, el meta-análisis se ha convertido en un importante complemento y a la vez reto de los métodos tradicionales de revisión de investigaciones (procedimientos narrativos, el método de la votación, etc.) demostrando su capacidad para inferir respuestas válidas a cuestiones fundamentales de las ciencias sociales. Su primera aplicación fue en la integración de estudios sobre los efectos de la psicoterapia (Smith, M.L., y Glass, G.V., 1977), que levantó una enorme polvareda de controversias, no acalladas en otras aplicaciones, de las que destacamos las realizadas sobre los efectos del «tamaño de la clase» (Glass, G.V., y Smith, M.L., 1979, Smith, M.L., y Glass, G.V., 1980).

No queremos ocultar en esta introducción los problemas iniciales que es necesario solventar si se quiere utilizar correctamente el meta-análisis. Entre los problemas señalamos: el *identificar y obtener los estudios más apropiados, extraer los estimadores del tamaño del efecto a partir de dicho estudios, codificar o clasificar los estudios, analizar los datos y publicar los resultados del análisis*. Finalmente diremos que la mejor fuente para quien desee documentarse técnicamente acerca del meta-análisis es, a nuestro juicio, el libro de Glass, G.V., McGraw, B., y Smith, M.L., (1981) «*Meta-analysis in Social research.*»

## FÓRMULAS ACUMULATIVAS MÁS UTILIZADAS

### A) TAMAÑO DEL EFECTO

El estadístico más utilizado en meta-análisis es el llamado «tamaño del efecto» y nombrado mediante una «d» (ya que no es otra cosa que la *diferencia entre las medias grupales divididas por la desviación típica*). Dicho estadístico ha sido popularizado por Glass y sus asociados (1978), podemos considerar dos variantes dentro del mismo, dependiendo de como estimemos la desviación típica, así Glass, G.V., y Smith, M.L., (1977) utilizan preferentemente como estimador la desviación típica del grupo control, mientras que Hunter, J.E., Schmidt, M.L., y Jackson, G.B., (1983) se inclinan por la desviación típica intra-grupo. (argumentan la bondad del segundo frente al primero a que se incurre en menos error de muestreo).

Definiremos pues el estadístico «d» como:

$$d = \frac{\bar{Y}_E - \bar{Y}_C}{S}$$

en el primer caso a que aludíamos anteriormente identificaremos:  $S = S_C$ ,

mientras que en el segundo:

$$S = \sqrt{\frac{(N_E - 1) S_E^2 + (N_C - 1) S_C^2}{N_E + N_C - 2}}$$

No acaban aquí las dificultades estadísticas que entraña el uso de estas fórmulas acumulativas, así entendemos que cualquier estadístico puede ser afectado por tres tipos de artefactos:

- errores de muestreo*
- errores de medición*
- errores por variaciones en el rango.*

Concretamente el estadístico «d» es fundamentalmente influído por el primero y accidentalmente por el segundo. Las técnicas meta-analíticas nos permiten controlar dichos efectos.

Así, y siguiendo a Hedges, L.V. (1982), podremos encontrar la media ponderada de las d y su varianza corregida de los posibles errores de muestreo. Las fórmulas que propone son las siguientes:

$$\bar{d} = \frac{\sum (N_i d_i)}{\sum N_i}$$

$$\sigma_d^2 = \frac{\sum [N_i (d_i - \bar{d})^2]}{\sum N_i}$$

$$\sigma_e^2 = \frac{4(1+\bar{d}/8)k}{N}$$

$$\sigma_d^2 = \sigma_d^2 - \sigma_e^2$$

Siendo: K el número de estudios

N el tamaño total de la muestra entre todos los estudios.

$N_i$  tamaño de la muestra de cada estudio i.

Si el tamaño del efecto es el mismo a través de todos los estudios entonces  $\sigma_d$  será aproximadamente igual a cero. Si la variación es grande, especialmente si lo es en relación al valor medio de la d, no nos quedará más remedio que buscar posibles *variables moderadoras* que la expliquen.

Si conocemos la fiabilidad de cada una de las pruebas empleadas en la medición de las variables dependientes podamos por atenuación, corregir los efectos de los errores de medición.

La fórmula a utilizar sería la siguiente:

*Para la d:*

$$d_{corregida} = \frac{d}{\sqrt{r_{yy}}}$$

Siendo:

$d_{corregida}$ : la nueva d

d: la anterior

$r_{yy}$ : el coeficiente de fiabilidad de la variable dependiente.

*Para la varianza del error* (también se ve afectada por errores de medición)

$$\sigma_e^2 = \frac{1}{\bar{r}_{yy}} \cdot \frac{4(1 + \frac{\bar{d}^2}{8}) K}{N}$$

El proceso general a seguir sería por tanto muy sencillo, en primer lugar obtendríamos las d (en el caso de poseer información acerca de la fiabilidad de las pruebas empleadas deberíamos corregirlas por atenuación). Con las d obtenidas procederíamos a buscar su media y desviación típica incial, que sería corregida posteriormente extrayéndole la varianza debida al error (que a su vez puede también ser modificada en el caso, anteriormente contemplado, de conocer datos acerca de la fiabilidad de las pruebas).

## B) ESTUDIOS CORRELACIONALES

Los estudios correlacionales mantienen la primacía en cuanto a estar entre los más empleados en la investigación educativa. De ahí, la importancia de dedicarles amplia atención en esta revisión de técnicas meta-analíticas.

El coeficiente de correlación se ve afectado también por los tres artefactos a que nos referíamos en el anterior apartado: *errores de muestreo, errores de medición y errores por variaciones en el rango*. El meta-análisis nos proveerá de medios para intentar salvarlos.

Si tenemos un conjunto de correlaciones correspondientes a diversos estudios, el mejor estimador de la correlación de la población será:

$$\bar{r} = \frac{\sum(N_i r_i)}{\sum N_i}$$

siendo:

$r_i$  = la correlación correspondiente al estudio i

$N_i$  = número de sujetos del estudio i

y en el de su varianza:

$$S_r^2 = \frac{\sum [N_i (r_i - \bar{r})^2]}{\sum N_i}$$

### 1) Errores de muestreo

Los errores de muestreo no afectarán a la correlación de la población pero sí a su varianza. Se impone por tanto su corrección.

Consideraremos que la verdadera varianza es:

$$\sigma_p^2 = \sigma_r^2 - \sigma_e^2$$

siendo:

$$\sigma_e^2 = \frac{(1 - \bar{r}^2)^2 K}{N}$$

K = número de estudios implicados en el meta-análisis  
 $N = \sum N_i \rightarrow$  tamaño total de la muestra

por tanto:

$$\sigma_p^2 = S_r^2 - \frac{(1 - \bar{r}^2)^2 K}{N}$$

## 2) Errores de medición

En el supuesto que conociéramos la fiabilidad de todas las pruebas utilizadas en cada uno de los estudios, podemos, por attenuación, corregir cada una de las correlaciones observadas.

$$r \text{ corregida} = \alpha \cdot r \text{ observada}$$

$$\text{Siendo } \alpha = \frac{1}{\sqrt{r_{xx} \cdot r_{yy}}}$$

La falta de fiabilidad también afectará a la varianza del error, que debemos, consiguientemente, corregir.

$$\sigma^2 \text{ corregida} = \alpha^2 \cdot \sigma^2 \text{ error}$$

## 3) Variaciones en el rango

El coeficiente de correlación se ve también afectado cuando los valores de la variable independiente difieren significativamente en los diferentes estudios. De hecho sólo serían comparables aquellas correlaciones obtenidas sobre poblaciones con la misma desviación típica en la variable independiente.

Existen fórmulas correctoras que nos permiten averiguar en qué medida la correlación ha podido ser afectada por la restricción (o ampliación) del rango.

Precisaremos definir primero el valor  $u$ , que definiremos como:

$$u = \frac{s}{S}$$

siendo  $s$ : desviación típica de la población del estudio

siendo  $S$ : desviación típica de la población de referencia

Si al pasar de la población de referencia a la del estudio, hay restricción en el rango « $u$ » tendrá un valor inferior a la unidad, por el contrario si ha habido ampliación será mayor que la unidad.

Cualquier correlación de la población referencial se vería alterada en el estudio concreto donde se ha producido la variación en el rango en la siguiente medida:

$$r \text{ estudio} = \frac{u \cdot r \text{ referencial}}{\sqrt{(u^2 - 1) r^2 \text{ referencial} + 1}}$$

Si por el contrario nosotros tenemos la correlación de un estudio específico y creemos que de alguna manera no refleja la verdadera correlación (ya que ha habido una drástica reducción en el rango) y preferimos corregirla de forma que transformemos aproximándola a la que hubiera tenido en la población de referencia, el proceso sería el siguiente:

$$U = \frac{1}{u}$$

$$r_{\text{referencial}} = \frac{U \cdot r_{\text{estudio}}}{\sqrt{(U^2 - 1) r_{\text{estudio}}^2 + 1}}$$

Quizás un pequeño ejemplo nos ayudaría a clarificar este último punto.

Supongamos que en un estudio realizado con alumnos de 8.<sup>º</sup> de EGB, se ha comprobado que la correlación entre un test de inteligencia y el rendimiento en Matemáticas es de 0,70. Al cabo de un año se vuelve a recoger información acerca de estos muchachos en 1.<sup>º</sup> de BUP, y se comprueba que la nueva correlación entre las mismas variables es de 0,44.

¿Ha variado sustancialmente el peso del factor inteligencia en el rendimiento matemático?

Se sabe que  $u = 0,50$

$$U = \frac{1}{0,5} = 2$$

$$r_{\text{referencia}} = \frac{2 \cdot 0,44}{\sqrt{(2^2 - 1) \cdot 0,44^2 + 1}} = 0,70$$

Se comprueba que una vez corregida la correlación, la diferencia observada es atribuible exclusivamente a la variación en el rango, no a una verdadera diferencia entre ellas.

Hasta ahora hemos supuesto que conocíamos toda la información necesaria para corregir correlación por correlación las alteraciones sufridas bien sea por errores de medición o por variaciones en el rango, para pasar posteriormente a buscar la media y la varianza de las correlaciones de todos los estudios y extraer finalmente la varianza del error atribuible al muestreo.

Sin embargo, lo habitual es no poder disponer de toda la información que nos permitiera actuar así. Aún en estos casos el meta-análisis provee de las técnicas necesarias para poder de alguna forma corregir los efectos de dichos arte-

factos. Procederemos justamente al revés de como lo habíamos hecho hasta ahora, primero calcularemos la media y la varianza de las correlaciones no corregidas para corregir posteriormente los efectos a que antes aludímos y que evidentemente no deseamos (las fórmulas para ello fueron desarrolladas por Schmidt, M.L., y Hunter, J.E., 1977).

Cabe considerar dos situaciones, la más común en que se observa errores de medición en ambas variables ( $x$  e  $y$ ) pero no restricción en el rango, y en las que intervienen los tres artefactos mencionados.

En el primer caso y utilizando una notación similar a la propuesta por Callender, J.C., y Osburn, H.G., (1980), procederíamos de la siguiente manera:

Llamaríamos:

$$a = \sqrt{r_{xx}}$$

$$b = \sqrt{r_{yy}}$$

Recogeremos estudios con estimaciones de  $r_{xx}$ ,  $r_{yy}$  y  $r_{xy}$  y tras las raíces (donde corresponde) buscaremos sus distribuciones y a partir de ellas obtendremos sus medias y sus varianzas.

Estimaremos los valores de la correlación de la población, a partir de:

$$\bar{\rho}_{xy} = \bar{r}_{xy}$$

$$\sigma_{\rho_{xy}}^2 = \sigma_{r_{xy}}^2 - \sigma_{\bar{r}}^2$$

Finalmente eliminaremos los efectos de los errores de medición y obtendremos los parámetros de la población ya definitivamente corregidos:

$$\bar{P} \text{ corregida} = \frac{\bar{\rho}_{xy}}{\bar{a} \bar{b}}$$

$$\sigma_{\bar{\rho}}^2 \text{ corregida} = \frac{\sigma_{\rho_{xy}}^2 - \bar{\rho}^2 \text{ corregida} (\bar{a}^2 \sigma_b^2 + \bar{b}^2 \sigma_a^2)}{\bar{a}^2 \bar{b}^2}$$

Seguiremos para el segundo caso con las notaciones de Callender y Osburn. Definimos los siguientes valores:

$$a = \sqrt{r_{xx}}$$

$$b = \sqrt{r_{yy}}$$

$$c = \frac{u}{\sqrt{(u^2 - 1) \rho^2 + 1}}$$

donde:  $u = \frac{\sigma_{\text{de la población del estudio}}}{\sigma_{\text{de la población referencial}}}$

También se puede averiguar «c» con una fórmula alternativa derivada por Callender, J.C., y Osburn, M.G., (1980).

$$c = \sqrt{u^2 + (1 - u^2) \rho_{xy}^2}$$

$$r_{xy} = \rho_{xy} + e$$

donde:

$$\rho_{xy} = a \cdot b \cdot c \cdot \rho_{\text{ref.}}$$

$\rho_{xy}$  = es la correlación de la población en el estudio.

$\rho_{\text{ref.}}$  = es la correlación en la población referencial.

$$p^2 = r_{xx} \cdot r_{yy} \cdot \rho_{\text{ref.}}^2$$

$$\bar{\rho}_{xy} = \bar{r}_{xy}$$

$$\sigma_{\rho_{xy}}^2 = \sigma_{r_{xy}}^2 - \sigma_e^2$$

$$\sigma_e^2 = \frac{(1 - \bar{r}_{xy}^2) K}{N}$$

Una vez definidos todos los valores podemos pasar a considerar las fórmulas definitivas, estas serán:

$$\bar{\rho}_{\text{Ref.}} = \frac{\bar{r}_{xy}}{\bar{a} \cdot \bar{b} \cdot \bar{c}}$$

$$\sigma_{\rho_{\text{ref.}}}^2 = \frac{\sigma_{\rho_{xy}}^2 - \bar{\rho}_{\text{Ref.}}^2 (\bar{b}^2 \cdot \bar{c}^2 \cdot \sigma_a^2 + \bar{a}^2 \cdot \bar{c}^2 \cdot \sigma_b^2 + \bar{a}^2 \cdot \bar{b}^2 \cdot \sigma_c^2)}{\bar{a}^2 \cdot \bar{b}^2 \cdot \bar{c}^2}$$

Dada la complejidad de las últimas formulaciones, incluyo a modo de ilustración un ejemplo extraído del libro de Hunter, J.E., Schmidt, F.H., y Jackson, G.B., (1983, pág. 87).

N	r <sub>xx</sub>	r <sub>yy</sub>	u	r <sub>xy</sub>	a	b	c
68	.49	—	.40	.02	.70	—	.43
68	—	.64	—	.26	—	.80	—
68	.49	.64	—	.33	.70	.80	—
68	—	—	.60	.09	—	—	.62
68	.49	—	—	.02	.70	—	—
68	—	.49	.40	.24	—	.70	.43
68	.49	.49	—	.30	.70	.70	—
68	—	—	.60	.06	—	—	.62
68	.64	—	.40	.28	.80	—	.43
68	—	.64	—	.04	—	.80	—
68	.64	.64	—	.12	.80	.80	—
68	—	—	.60	.34	—	—	.62
68	.64	—	—	.26	.80	—	—
68	—	.49	.40	.02	—	.70	.43
68	.64	.49	—	.09	.80	.70	—
68	—	—	.60	.33	—	—	.62

Las cuatro medias y varianzas que necesitamos son:

$$\bar{r}_{xy} = 0,175$$

$$\bar{a} = 0,75$$

$$\bar{b} = 0,75$$

$$\bar{c} = 0,525$$

$$\sigma_r^2 = 0,0151$$

$$\sigma_a^2 = 0,0025$$

$$\sigma_b^2 = 0,0025$$

$$\sigma_c^2 = 0,009$$

Usamos la media y la varianza de las correlaciones observadas para corregir los errores de muestreo:

$$\bar{\rho}_{xy} = \bar{r}_{xy} = 0,175$$

$$\sigma_{\rho_{xy}}^2 = \sigma_{r_{xy}}^2 - \sigma_E^2 = 0,0151 - 0,013819 = 0,001281$$

Finalmente pasamos a corregir los valores, eliminando los efectos de los errores de medición y de variación en el rango:

$$\bar{\rho}_{Ref} = \frac{0,175}{(0,75)(0,75)(0,525)} = 0,59$$

$$\sigma_{\rho_{Ref}}^2 = \frac{0,001281 - 0,59^2 (0,75^2)(0,525^2) 0,0025 + 0,75^2 (0,525^2) \cdot 0,0025 + 0,75^2 (0,75^2) 0,009}{(0,75^2)(0,75^2)(0,525^2)} = 0,000229$$

Es decir que la verdadera correlación de la población podemos cifrarla en 0,59 y  $\sigma_p = 0,015$ , muy cercano a  $\bar{\rho} = 0,60$  y  $\sigma_p = 0,00$  que fueron los valores utilizados para construir la tabla.

## SÍNTESIS

A fin de intentar no perder una visión global de tema, perdidos en tantos detalles, resumiríamos esta parte de revisión metodológica indicando que cualquier investigación está sometida, a tres fuentes de variación que podemos controlar mediante el meta-análisis: *Errores de muestreo, errores de medición y variaciones en el rango*.

Los dos primeros son habituales en cualquier estudio y un meta-análisis completo debería corregir ambos, exceptuando naturalmente el caso de que no sean publicados los datos referentes a la fiabilidad de las medidas usadas. Hay algunas áreas de investigación que habrá que añadir el control sobre los efectos derivados de las restricciones en el rango.

Simplemente conociendo los tamaños de las muestras  $N_i$  y cada correlación  $r_{ij}$  se pueden corregir los efectos de los errores de muestreo. La media y la varianza de las correlaciones observadas en las muestras son usadas para estimar la media y la varianza de las correlaciones de la población.

Tras corregir los errores de muestreo podemos enfrentarnos a eliminar los errores de medición, dándose habitualmente dos posibilidades: que tengamos información de la fiabilidad para cada estudio o solamente de algunos. En el primer caso se corrige cada correlación por separado por atenuación y se procede a meta-analizarlas posteriormente. En el segundo se procede primero a corregir los errores de muestreo de la varianza de las correlaciones observadas para pasar posteriormente a corregir los efectos de los errores de medición mediante la información que tengamos acerca de la distribución de las fiabilidades.

En el caso que nos enfrentemos a los tres artefactos se procede de forma análoga al apartado anterior. Si se tiene toda la información se corrige primero los errores de medición y los debidos a variaciones en el rango para proceder posteriormente a efectuar el meta-análisis. En el caso de que no se posea la información completa deberemos como antes tras eliminar los errores de muestreo servirnos de las fórmulas generadas por las distribuciones para corregir los otros dos efectos.

Si la varianza una vez corregida difiere significativamente de cero puede estar indicando la presencia de variables moderadoras, que deberemos localizar bien sea analizando los estudios clasificados por subconjuntos o bien categorizándolos y generando vectores susceptibles de ser analizados mediante la técnica de la regresión.

Finalmente creo que es importante el señalar que a pesar de la popularidad del estadístico  $d$ , personalmente me inclino más a sustituirlo por un índice de

correlación. En el fondo el estadístico  $d$  no es sino una transformación de la correlación biserial puntual (ver Glass, McGaw y Smith (1981)).

$$r_{bp} = \frac{d}{\sqrt{d^2 + 4(N-2)}} \\ N$$

Con la ventaja de que una correlación biserial puntual pueda ser insertada en una matriz de correlaciones y pueda ser tratada métricamente a través de diversas pruebas (correlación parcial, análisis de sendas, etc. etc.). Podemos aplicar todo lo dicho hasta aquí de la correlación de Pearson a la biserial puntual, simplemente teniendo en cuenta que ésta necesita ser corregida en el caso de que los tamaños de las muestras sean distintos:

$$r \text{ corregida} = \frac{r_{bp}}{\sqrt{4pq(1-r_{bp}) + r_{bp}^2}}$$

$$\text{siendo } pq = \frac{N_e N_c}{N^2}$$

## UNA APLICACIÓN ESPECÍFICA

Vamos a exponer a continuación un meta-análisis llevado a cabo con veinte estudios correlacionales, que tratan (muchos de ellos no exclusivamente) sobre la aportación de las variables aptitudinales en la explicación de la varianza del rendimiento escolar. Dichos estudios corresponden a los últimos años (1977-85) y aunque no está la total población de posibles trabajos que traten sobre el tema, si constituyen, a nuestro juicio, una muestra representativa de los realizados últimamente en España.

Tal como señalamos en un anterior trabajo (Garanto, Mateo y Rodríguez 1985), la relación entre factores aptitudinales y rendimiento escolar presenta una constancia en los resultados que no se da con otras variables (de personalidad por ejemplo). Sin embargo este hecho intuído por muchos no ha conducido a una constatación objetivo-estadística que nos permita zanjar el tema y sigamos incluyendo en cualquier estudio correlacional sobre el tema las mismas variables y utilizando idéntica metodología, que nos conducirá irremediablemente a parecidos resultados, gastando energía inútilmente en algo que quizás ya deberíamos considerar superado.

Presentamos a continuación un cuadro resumen con los 20 estudios mencionados anteriormente, para posteriormente meta-analizarlos:

CUADRO RESUMEN DE ESTUDIOS

AUTOR	NIVEL	MUESTRA	TIPO PREDICTORES	N.º	CRITERIO RENDTO.	R	R <sup>2</sup>
Pelechano, V. (1977)	1.º BUP	2.487	Intelec. (PMV-V, PMA-R EVR Analog. y EVR Inferenc.)	4	Global curso	.317	.1005
Pérez, M. (1979)	5.º EGB	470	Intelec. (THE, THG, TCV)	3	Global curso	.364	.1325
Martínez, A. (1980)	Alumnos: E.U. Form. Profesor.	181	Intelec. (Puntuación conjunta varias pruebas)	1	Global materias comunes.	.470	.2209
Pérez Gzlez. J. (1980)	1.º EGB	208	Intelec. (ABC)	8	Expr. Verbal	.389	.1513
Pérez Gzlez. J. (1980)	1.º EGB	208	Intelec. (ABC)	8	Expr. Numérica	.386	.1490
Pérez Gzlez. J. (1980)	1.º EGB	208	Intelec. (Aptitudes Cognoscitivas de Thorndike, Hagan y Lorge).	4	Expr. Verbal	.419	.1756
Pérez Gzlez. J. (1980)	1.º EGB	208	Intelec. (Aptitudes cognoscitivas de Thorndike, Hagan y Lorge).	4	Expr. Numérica	.479	.2294
Corominas, E. (1981)	8.º EGB	385	Intelec. (D-48, PMA-Vyy, PMA-N)	3	Media Leng. Mat. y A. Social	.570	.3249
Corominas, E. (1981)	8.º EGB	285	Intelec. (D-48, PMA-Vyy, PMA-N)	3	Matemáticas	.532	.2820
Estany Bassa, S. (1981)	FP 1.º	114	Intelec. (D-70, PMA-RyV, García Yague APT-N)	5	Global curso	.540	.2916
Estany Bassa, S. (1981)	FP 2.º	42	Intelec. (D-70, PMA-RyV, García Yague APT-N)	5	Global curso	.380	.1444
Seisdedos, N. y López, N. (1981)	8.º EGB	246	Intelec. (Factores: Ap. Técnicas, Numérica y Verbal)	3	Global curso	.382	.1459
Rodríguez, S. (1982)	6.º EGB	340	Intelec. (D-48, K-A, DAT-AR)	3	Media Leng. Mat. y A. Social	.503	.2530
Rodríguez, S. (1982)	7.º EGB	326	Intelec. (D-48, K-A, DAT-AR)	3	Media Leng. Mat. y A. Social	.378	.1429
Rodríguez (1982)	8.º EGB	334	Intelec. (D-48, K-A, DAT-AR)	3	Media Leng. Mat. y A. Social	.406	.1648
García, J. (1983)	8.º EGB	262	Intelec. (D-48, DAT-VR y DAT-NA)	3	Medio Leng. Mat. y A. Social	.474	.2246
García, J. (1983)	8.º EGB	262	Intelec. (D-48, DAT-VR y DAT-NA)	3	Matemáticas	.438	.1918
Turón, J. (1984)	1.º Univ. Biológ.	103	Intelec. (D-48, DAT-VR y DAT-NA)	3	Matemáticas	.430	.1849
Buendia Eisman, L. (1985)	Ciclo Inc. y Medio	1190	Intelec. (Test Aptitud. escolares TEA)	5	Nota Global	.383	.1467
Garanto, J. Mateo, J. y Rodríguez, S. (1985)	8.º EGB	251	Intelec. (Escala colectiva de Nivel. Intelect. de Pierre Benedetto)	5	Media Leng. y Matemáticas	.532	.2830

## Procedimiento de cálculo:

$N_i$	$r_i$	$N_i r_i$	$r_i - r$	$(r_i - r)^2$	$N_i (r_i - r)^2$
2487	0,317	788,379	0,083	0,006889	17,133
470	0,364	171,08	0,036	0,001296	0,609
208	0,389	80,912	0,011	0,000121	0,025
208	0,386	80,288	0,014	0,000106	0,041
208	0,419	87,152	0,019	0,000361	0,075
208	0,479	99,632	0,079	0,006241	0,003
181	0,470	85,07	0,07	0,0046	0,887
385	0,570	219,45	0,17	0,0289	11,126
385	0,532	204,82	0,132	0,017424	6,708
114	0,540	61,56	0,14	0,0196	2,234
42	0,380	15,96	0,02	0,0004	0,017
246	0,382	93,972	0,018	0,000324	0,080
340	0,503	171,02	0,103	0,010609	3,607
326	0,378	123,228	0,022	0,000484	0,158
334	0,406	135,604	0,006	0,000036	0,012
262	0,474	124,188	0,074	0,005476	1,435
262	0,438	114,756	0,038	0,001444	0,278
103	0,430	44,29	0,03	0,0009	0,093
1190	0,383	455,77	0,017	0,000289	0,344
251	0,532	133,532	0,132	0,017424	4,373

 $\Sigma: 8210$  $\Sigma: 3290,663$  $\Sigma: 49,338$ 

$$\bar{r} = \frac{3290,663}{8210} = 0,4008$$

$$S_r^2 = \frac{49,338}{8210} = 0,0060095$$

$$\sigma_e^2 = \frac{(1 - 0,4^2) \cdot 20}{8210} = 0,002$$

$$\sigma_p^2 = 0,006 - 0,002 = 0,004$$

$$\sigma_p = 0,06$$

Hemos incluido únicamente a modo de ilustración los cálculos manuales del meta-análisis. En realidad fueron averiguados (y nos sirvieron de comprobación) con una sencilla calculadora que incluía algunas funciones estadísticas (Media y desviación típica).

La correlación de la población podemos fijarla en 0,40 (pensamos para valorar dicho índice, que la muestra de estudios incluye a 8210 sujetos). El valor de la desviación típica 0,06, es lo suficientemente bajo como para que no sea necesario el pensar en identificar variables moderadoras.

Finalmente cabe añadir que al no disponer de más información no hemos podido eliminar los otros artefactos mencionados en nuestro trabajo. Proponemos que en las publicaciones se exija el incluir al menos aquellos índices que facilitarán este tipo de análisis.

#### REFERENCIAS BIBLIOGRÁFICAS

- BUENDIA, L., (1985). *Factores Determinantes del Rendimiento en E.G.B.* I.C.E. Universidad de Granada.
- CALLENDER, J.C. y OSBURN, H.G., (1980) Development and test of a new model for generalization of validity. *Journal of applied Psychology*. 65, 543-558.
- COROMINES, E. (1981) *Predicción del rendimiento en alumnos de 8.º de E.G.B. Estudio longitudinal*. Tesina de licenciatura inédita. Universidad de Barcelona.
- ESTANY, S. (1981) *Batería Predictiva i Diferencial per a FP 1*. Tesis de Licenciatura inèdita. Universitat de Barcelona.
- GARANTO, J., MATEO y RODRÍGUEZ, S. (1985). Modelos y técnicas de análisis del rendimiento académico. En prensa. Barcelona.
- GARCÍA, J., (1983) *La motivación escolar en un modelo predictivo del rendimiento*. Tesis de licenciatura inédita. Universidad de Barcelona.
- GLASS, G.V. (1976). Primary, secondary, and meta-analysis of research. *Educational Research*, 5, 3-8.
- GLASS, G.V., (1978) Integrating findings: The meta-analysis of research. In. L.S. Shalman (Ed.) *Review of Research in Education*, 5, F.E. Peacock, Itasca.
- GLASS, G.V., (1980). On criticism of class size/student achievement research: No points conceded. *Phi Delta Kappa*. 242-244.
- GLASS, G.V., McGAW, B. y SMITH, M.L., (1981). *Meta-analysis in social research*. Sage. Beverly Hills..

- GLASS, G.V., y SMITH, M.L., (1979) Meta-analysis of the relationship between class-size and achievement. *Educational Evaluation and Policy Studies*, 1, 2-16.
- HEDGES, L.V. (1983) *Statiscal Methodology in Meta-analysis*. Eric/T.M. Report 83.
- HUNTER, J.E., SCHMIDT, F.L., y JACKSON, G.B., (1983) *Meta-Analysis. Cumalating Research findings across studies*. Sage, Beverly Hills.
- MARTÍNEZ, A., (1980) Estudio analítico del rendimiento académico de los grupos. *Rev. Española de Pedagogía*, 145, 29-41.
- PELECHANO, V., (1977) *Personalidad, inteligencia, motivación y rendimiento en el B.U.P.* I.C.E. de la Universidad de La Laguna.
- PÉREZ GONZÁLEZ, J., (1980). Predictores en los inicios del aprendizaje formal. *Rev. Española de Pedagogía*, 149, 113-120.
- PÉREZ, M., (1979). Relación entre fracaso escolar, timidez y rendimiento. *Revista de Educación*, 99-10, 299-337.
- RODRÍGUEZ, S., (1982) *Factores de rendimiento escolar*. Oikos-Tau. Barna.
- SEISDEDOS, N. y LÓPEZ, N., (1981). Las evaluaciones y una batería de tests (en 8.º EGB). *Rev. de Psicol. Gral. y Aplicada*, 169, 255-262.
- SMITH, M.L., y GLASS, G.V., (1977) Meta-Analysis of psychotherapy outcome studies. *American Psychologist*, 32, 752-760.
- SMITH, M.L., y GLASS, G.V., (1980). Meta-analysis of class size and its relaationship to attitudes and instruction. *American Educational Research Journal*, 17, 419-433.
- SCHMIDT, F.L., y HUNTER, J.E., (1977) Development of a general solution to the problem of validity generalization. *Journal of Applied Psychology*, 62, 529-540.
- THORNDIKE, R.L., (1933) The effect of the internal between test and retest on the constancy of the I.Q. *Journal of Educational Psychology*, 25, 543-549.
- TURÓN, J., (1984). *Factores de rendimiento académico en la Universidad*. EUNSA. Pamplona.

## Estrategias metodológicas en los estudios longitudinales

por  
*Mario de Miguel*

### RENDIMIENTO ACADÉMICO Y PROCESOS DE ESCOLARIZACIÓN

Uno de los tópicos que ha centralizado mayor número de trabajos durante las últimas décadas ha sido la evaluación de la eficacia de los procesos de escolarización (School effectiveness). Basta recordar algunos estudios (Coleman, J., 1986, Plowden, B., 1966, Húsen, T., 1967, Jencks, C., y col. 1972, Comber, L. y Keeves, J. 1973...) para captar la importancia que ha tenido este tipo de trabajos en el desarrollo de la metodología de la investigación educativa. También es conocido que los resultados iniciales de estos trabajos han originado un evidente pesimismo acerca del papel que juegan las escuelas en el contexto social, al demostrar el poco impacto que estas tienen en el desarrollo personal de sus alumnos. Como ya hemos señalado anteriormente (De Miguel, M., 1984), los resultados de estos estudios, reforzados por las evaluaciones negativas de los primeros programas compensatorios, han supuesto una auténtica losa para cuantos consideraban la escuela como medio para paliar las diferencias socioculturales y ofrecer a todos los alumnos una auténtica igualdad de oportunidades.

La situación actual es ciertamente distinta. De una parte, la diversidad de juicios sobre qué se entiende por eficacia de una escuela impide establecer un criterio de referencia válido que pueda utilizarse de forma generalizada como variable output a corto o largo plazo (Astin, A., 1977, Pace, C., 1979, Bridge, R., Judd, C., y Moock, P., 1979, Madaus, G., Airasian, P., y Kellaghan, T., 1984). Aunque existe la convicción de que las escuelas no facilitan a todos los niños las máximas oportunidades para su desarrollo personal, no se puede extraer/polar/generalizar sobre la ineficacia de los procesos de escolarización cuando no existe aún acuerdo sobre los fines que estas instituciones deben cubrir dentro del orden social. La diversidad de opiniones al respecto hace que el problema se plantee –en última instancia– en términos de valores, creencias y/o preferencias que relativizan cualquier diseño metodológico, incluso aquellos que se establecen en términos mercantilistas tomando como criterio de eficacia el correlato

entre lo que se proyecta hacer y lo que efectivamente se hace. La habitual utilización del rendimiento académico del alumno como criterio de las variables output hace aún más difícil aislar el efecto de la escuela, dado que esta variable está estrechamente relacionada con las habilidades intelectuales y el entorno social del sujeto en una constante y permanente interacción. Mientras que desde la investigación no se clarifiquen las diferencias asociadas con las influencias extraescolares (background), difícilmente se podrán aislar criterios válidos de rendimiento dependientes exclusivamente de la escuela, dado que ambas variables actúan asociadas (*in consort*).

De otro lado, la revisión de las investigaciones iniciales sobre esta temática (Smith, M., 1972, Mayeske, G., y col. 1972, Mosteller, F., y Moynihan, D., 1972, Mayeske, G., y col. 1973), ha permitido detectar una serie de fallos metodológicos en los trabajos que anulan, en buena parte, los resultados obtenidos. Entre ellos cabe destacar los siguientes:

- Limitaciones de los paradigmas de investigación utilizados para describir los efectos de la escolarización como distintos a los procesos específicos de maduración, aprendizaje y curriculum (Romberg, T.H., 1981), así como de otros agentes educativos.
- Dificultades para representar la complejidad de los factores interviniéntes en la escolarización en términos de variables input, particularmente si tales variables deben ser operativizadas en términos estadísticos (Averch, H., y col. 1972, Kellaghan, T., 1977).
- Dificultades para controlar e incluir en el diseño la diversidad de interacciones que se producen dentro de la escuela así como de las que se establecen entre la escuela y el entorno socio-familiar, máxima cuando presentan mayor efectividad aquellas variables relativas a la actividad (proceso) de la escuela frente a las que solamente tipifican su status (Marjoribanks, E., 1974, Good, T., Biddle, J., y Brophy, J., 1975, Crawford, J., y col. 1977, Bloom, B., 1976).
- Limitaciones derivados de la escasez y poca adecuación y sensibilidad de los instrumentos de medida para captar los efectos derivados de los factores de escolarización (Rosenshine, B., y Furst, N., 1971, Cooley, W., 1974, Klein, S., 1976, Brookover, W., y col. 1978, Brimer, A., y col. 1978).
- Problemas de carácter metodológico subyacentes en los diversos diseños utilizados y otros relativos a la técnica de análisis –de sesgo, errores de medida, agrupación de las varianzas, contrastes adecuados, aditividad y colinearidad de variables, etc.– (Walberg, H., y Rasher, S., 1974, Bidwell, C., 1975, Wiley, D., 1976, Rakow, E., y col. 1978, Madaus, G., y col. 1980).

Todas estas limitaciones, originadas al tratar de ajustar/representar una realidad compleja en diseños simplificados, han determinado que muchas de las

conclusiones o afirmaciones de los estudios sobre la eficacia de las escuelas se hayan relativizado. Evidentemente, mientras que los criterios de evaluación no sean uniformes y los diseños adolezcan de limitaciones conceptuales y metodológicas ciertamente resulta arriesgado descalificar todo el sistema escolar como ineficaz. Por ello, más que en la ineficacia de las escuelas debemos pensar en las limitaciones de nuestros modelos y técnicas de análisis de la realidad escolar. Aunque han sido evidentes los avances metodológicos experimentados durante las dos últimas décadas –a pesar de los escasos presupuestos destinados a este área de investigación– aún estamos lejos de poder captar con nuestras actuales redes/instrumentos de trabajo la compleja interacción entre los factores interviniéntes en los fenómenos educativos. Siempre cabe pensar que con un tratamiento más igualitario en relación con otras áreas científicas se podría esperar en un futuro próximo un avance espectacular en relación al tema que nos ocupa.

## METODOLOGÍA LONGITUDINAL Y EXPLICACIÓN CAUSAL.

Una de las limitaciones más serias implícita en la mayoría de los trabajos sobre la eficacia de las escuelas es que –normalmente– se realizan en base a diseños de tipo transversal (cross-sectional) relativos a un momento dado cuando por su propia naturaleza los efectos de la escolarización se manifiestan en el tiempo. Muchos autores han incidido reiteradamente sobre este problema aconsejando la utilización de *estudios longitudinales* –diseños que requieren, al menos, dos evaluaciones sobre los mismos sujetos a lo largo de su vida– dado que estos son más apropiados para efectuar una estimación directa del cambio intraindividual y de las diferencias interindividuales atribuibles a los procesos de escolarización.

La escasez de investigaciones longitudinales en el campo educativo es patente. En una revisión reciente de este tipo de estudios (Harway, M., y Mednick, S., 1984) se pone de manifiesto el bajo nivel de desarrollo alcanzado en comparación con otras áreas afines –psiquiatría, sociología, medicina,...– ya que tan sólo representan el 4% del total de trabajos recogidos lógicamente esta escasez de estudios longitudinales se relaciona con las dificultades específicas de carácter metodológico –selección de variables apropiadas, utilización de instrumentos sensibles a los fenómenos temporales, muestreo, seguimiento de los sujetos, repetición de medidas, fiabilidad de datos, etc.– y práctico –estabilidad del equipo investigador, costos, manejo de datos, publicaciones a largo plazo,...– que lleva la realización de este tipo de estudios (Harway, M., Mednick, S., 1984).

Sin embargo, a pesar de estas y otras limitaciones evidentes, en la actualidad este tipo de estudios está cobrando gran impulso en la investigación educativa dadas las indudables ventajas que aporta frente a los diseños transversales. A modo de síntesis se podrían resumir en estos puntos:

- Los estudios longitudinales permiten obtener/estimar el rendimiento de los alumnos antes y después de que hayan estado sujetos a un determinado periodo de escolarización, mientras que en los transversales no sólo no existe criterio referencial anterior sino que la medida de las aptitudes y la estimación de los factores del entorno (background) son tomadas al mismo tiempo que los resultados escolares.
- Los estudios longitudinales, dado que conllevan medidas repetidas sobre los mismos sujetos, posibilitan la utilización de estimaciones iniciales como línea-base a partir de la cual se puede medir/valorar el crecimiento o cambio individual así como una estimación de la incidencia de los factores de la escuela (tratamientos).
- Los estudios longitudinales no sólo juegan un papel importante en la observación de los cambios sino que al investigar las interconexiones de los factores determinantes del rendimiento ponen de relieve la magnitud y persistencia de estos factores en el tiempo (predicción), y posibilitan establecer, anticipadamente, programas de prevención.
- Los estudios longitudinales permiten superar algunas de las limitaciones metodológicas de los trabajos transversales al controlar los efectos de la maduración y las variables características de una época determinada e, incluso, cuando se establecen como modelos secuenciales posibilitan también aislar el efecto de las cohortes (variables generacionales).
- Finalmente, los estudios longitudinales, son especialmente importantes en orden a demostrar que determinadas variables (previsiblemente causas) preceden temporalmente a otras (posibles efectos), condición *sine qua non* que deben reunir todos los diseños que pretendan investigar la causalidad. Las correlaciones basadas en estudios transversales al no aportar tal evidencia no pueden ser utilizadas con tal fin.

Precisamente la importancia que cobran actualmente los estudios longitudinales tienen que ver con esta última finalidad: *la búsqueda de la explicación causal de los fenómenos educativos*. Todos sabemos que la simple constatación de que dos variables están asociadas estadísticamente –estudios correlacionales– no nos permite hablar de términos de causa-efecto (independiente-dependiente). Para que ello sea posible se deben cumplir, además, otros dos requisitos: *que la variable antecedente preceda temporalmente al consecuente (orden causal)* y *que tal asociación no desaparezca cuando se suprimen los efectos de otras variables precedentes (espureidad)*. Algunos autores (Hyman, H., 1963) recomiendan añadir otro criterio –que las variables causa-efecto estén relacionadas a través de una o varias intermedias– dado que es más útil para el investigador conocer el proceso a través del cual una variable influye en otra que saber solamente que es su causa.

En consecuencia, cualquier tipo de análisis correlacional que pretenda ir más allá de la simple constatación de una asociación debe especificar/garantizar el orden temporal entre las variables que se relacionan, estableciendo un modelo

o secuencia temporal donde se especifique en qué orden se manifiestan las variables que utilizamos. Por ello, antes de todo proceso de explicación-predicción de un fenómeno debemos proceder a discriminar las variables que entran en interacción con el mismo en las diferentes etapas de su desarrollo –*developmental sequence*– evitando no incurrir en la frecuente confusión de seleccionar las variables según criterios de proximidad (mayor correlación). En ciencias sociales las verdaderas relaciones causales suelen manifestarse en el tiempo, normalmente a través de otras variables intermedias, por lo que resulta imprescindible dicha especificación.

No obstante los problemas de la inferencia no acaban aquí. Además del orden causal debemos probar que las relaciones obtenidas son reales, no artificiales (espureidad). Es decir, demostrar que la relación entre nuestras variables no proviene de otras variables antecedentes. Normalmente para probar la artificialidad o no de una relación el investigador acude a otras hipótesis explicativas utilizando variables que preceden en el orden temporal a nuestra variable dependiente, aunque en la práctica –especialmente cuando se conocen los trabajos afines– se suele limitar a aquellas que le parecen más verosímiles. Lógicamente probar la no espureidad es ciertamente difícil dado que las posibilidades de variables alternativas son innumerables, máxime cuando, en ocasiones, los antecedentes (causas) derivan de los mismos procesos de investigación (contaminación). Por lo que el investigador siempre debe acabar la evaluación de sus trabajos en términos de la teoría de la falsación.

La importancia de los estudios longitudinales en relación al análisis causal es clara: *aportan al investigador una secuencia temporal de variables que pueden ser utilizadas como antecedentes y/o intermedias en una cadena causal*. Todos sabemos que una asociación estadística –fuerte o débil– no significa nada por sí misma, y que el investigador debe proceder a interpretar y/o explicar si estas relaciones son ciertas o simplemente aparentes. Los estudios longitudinales no sólo garantizan el segundo requisito o condición sino que también amplian el espectro de variables que pueden ser utilizadas para probar la no espureidad. Cuando la investigación tiene por finalidad la búsqueda de efectos a largo plazo –como en el caso de los trabajos sobre la eficacia de las escuelas– establecer esta secuencia de variables se hace imprescindible. De ahí el impulso que últimamente han tenido este tipo de estudios en la investigación del tema que nos ocupa (Weikart, A. 1970; Kraus, P., 1973; Astin, A., 1977; Entwistle, S., y Hayduc, L., 1979; Steel, L., Wise, L., y Abeles, R., 1979; Fletcher, J., Satz, P., y Morris, R., 1980).

## PRINCIPALES DISEÑOS EN LA INVESTIGACIÓN LONGITUDINAL

No existe una única texonomía de los estudios longitudinales dado que su clasificación puede realizarse en función de diversidad de criterios (Nesselroade, J., y Baltes, P.B., 1979, Mednick, S., Harway, M., y Finello, K., 1984). Aunque

las revisiones recientes aconsejan especificar unos datos mínimos a efectos de su tipificación/comparación posterior (tipo de variables observadas, edad, tamaño y tipo de cohorte estudiada, duración temporal de la investigación, criterios en la selección de la muestra, etc...), lo cierto es que, en la práctica, la mayor parte de las tipologías de estos estudios suele realizarse en función de dos criterios: a) según el tipo de diseño y b) en función del número de réplicas muestrales o cohortes utilizadas.

Desde la perspectiva del diseño metodológico los estudios longitudinales realizados en el campo educativo se pueden agrupar en dos grandes modelos: *estudios comparativos (ex post facto)* e *investigaciones cuasiexperimentales*. Las primeras parten de un modelo en el que consideran la investigación como un conjunto de estudios transversales donde cada uno de ellos es utilizado como replicación/criterio de comparación respecto a los anteriores, o si se prefiere donde la maduración o experiencias de la vida es el único tratamiento. En estos casos las técnicas de análisis se centran, fundamentalmente, en descubrir y relacionar los resultados obtenidos sobre un mismo grupo con otras medidas y/o factores antecedentes establecidos previamente (*investigación longitudinal correlacional*). Dentro de esta modalidad se incluyen todos los trabajos epidemiológicos, de seguimiento-crecimiento, retrospectivos y prospectivos, muy útiles en el campo educativo ya que permiten una estimación de los cambios desde una perspectiva temporal: elaboración de tipologías de problemas educativos según la edad, establecer comparaciones entre sexos, clases sociales y países, establecer asociaciones entre problemas de distintos tipos en momentos evolutivos determinados, demostrar la incidencia que tienen las variables del entorno y/o tratamiento educativo sobre los resultado escolares, etc... Los estudios longitudinales de carácter comparativo, a pesar de las limitaciones metodológicas de las que posteriormente hablaremos, cubren un objetivo básico en la investigación educativa: no sólo posibilitan una tipificación de los problemas escolares –*incidencia*– sino que permiten estimar su persistencia o remisión –*duración*–, al mismo tiempo que, al poner de manifiesto factores asociados, descubren predictores de futuros problemas (Robins, L., 1984).

El segundo modelo utilizado en los diseños longitudinales conceptualiza la investigación como una serie de *cuasiexperimentos* en los que el investigador trata de controlar/explicar la varianza que se produce entre dos mediciones sucesivas en función de determinadas variables experimentales que maneja. En estos casos la medición inicial es tomada como línea-base a partir de la cual se intenta dar una explicación a los cambios intra e inter individuales producidos por determinados tratamientos o programas de intervención educativa. En esta línea encontrariamos –dentro del tema que nos ocupa– todas las investigaciones realizadas para comprobar la eficacia de determinadas estrategias compensatorias con niños desaventajados socioculturalmente, que permiten conocer no sólo la eficacia de determinados tratamientos o programas escolares sino también, la plasticidad o modificabilidad de la conducta en las diferentes edades, y las posibilidades de optimizar/recuperar determinados defectos/problemas en el proceso

educativo. Las limitaciones subyacentes a la utilización de la metodología experimental en el campo educativo –máxime cuando se trata de diseños longitudinales– hace que sea muy limitado el número de trabajos realizados.

Respecto al segundo criterio –número de réplicas– los estudios longitudinales se subdividen en: *simples* y *secuenciales*. Los *estudios longitudinales simples* utilizan solamente un grupo muestral de observación –cohorte– sobre el que establecen varias mediciones a lo largo del tiempo. Tales estudios aunque suponen un avance sobre los transversales –donde se confunde el cambio individual (ontogenético) con las diferencias generacionales– también presentan una seria limitación: en ocasiones el cambio individual puede ser confundido con efectos de cambios socioculturales ocurridos durante el periodo temporal del estudio (Shaie, K., 1967, Baltes, P., B., Reese, H., y Nesselroade, J., 1979, Nesselroade, J., y Baltes, P.B., 1979, Shaie, K.Y. Hertzog, C., 1982). Ciertamente todos sabemos que existen acusadas diferencias entre las generaciones y los medios socio-culturales en el modo de concebir y tratar los problemas educativos, por lo que los estudios longitudinales simples presentan serias limitaciones relativas a su validez. Algunos autores sugieren como alternativa un diseño en el que se compara diferentes cohortes en una misma edad –*time-lag method*–, pero tal diseño también tiene su limitación: confunde las diferencias generacionales con los sucesos socioculturales ocurridos entre los dos períodos de evaluación.

Por ello, la única alternativa válida son los modelos en base a *estrategias secuenciales* en los que el investigador utiliza dos o más réplicas a fin de controlar las limitaciones de los estudios simples. El más sencillo es el diseño de *cohortes-secuencial* en el que se estudian dos o más cohortes sucesivas en un mismo espacio temporal con el fin de distinguir/separar el cambio individual de aquel que es debido a una cohorte específica. Si lo que se pretende es aislar las diferencias generacionales de las diferencias debidas a la edad entonces se parte de un *diseño secuencial-temporal* en el que muestras generacionales independientes son controladas en el mismo periodo de edad en diferentes tiempos. Finalmente encontrariamos los *diseños transversales-secuenciales* –dos o más cohortes sucesivas son examinados en dos o más momentos– que nos permita separar los efectos atribuibles a las diferencias generacionales y aquellos otros debidos al cambio sociocultural específico de un momento/tiempo dado.

Como se puede ver, la finalidad básica de estos modelos es controlar los *componentes fundamentales de la varianza temporal*: la maduración (edad), características generacionales (cohorte) y los sucesos/acontecimientos socioculturales propios de una época (tiempo de medición). La elección de una u otra estrategia depende de los planteamientos teóricos que el investigador hace sobre las relaciones entre las variables. Su aplicación en investigaciones relativas a los efectos de la escolarización no necesita mayor justificación dado que, por su propia naturaleza, el producto de la educación/réndimiento académico puede ser concebido y/o explicado, en buena parte, a partir de las tres variables anteriores.

Sin embargo la contribución más importante de las estrategias secuenciales a la investigación educativa es que han permitido –desde la metodología correla-

cional— *elaborar modelos con fines explicativos*. Como ya hemos dicho anteriormente, dado que en los estudios longitudinales quedan perfectamente delimitadas las variables en el orden temporal —antecedentes y consecuentes— es posible utilizar técnicas de análisis correlacional en base a paneles de intervalos cruzados, —cross-lagged— en los casos univariados, y path analysis y análisis estructural en los casos multivariados, para abordar modelos causales desde datos descriptivos. Hacia esta línea apuntan las orientaciones recientes de los principales autores que tratan de superar las limitaciones de las dos vías clásicas en la metodología de la investigación (Jöreskog, K., 1978, Rogosa, D., 1979, Kenny, D.R., 1979, Cox, T.H., y Campbell, D., 1979, Schaie, K., y Hertzog, C., 1982, Bentler, P., 1983).

Los modelos correlacionales a base de *paneles de intervalos cruzados (cross-lagged)* se utilizan exclusivamente en el contexto de los estudios longitudinales y parten de un diseño en el que especifican y analizan las relaciones entre dos variables en dos o más momentos distintos bajo el supuesto de que una diferencia significativa entre las correlaciones cruzadas ( $r_{x_1y_2}, r_{y_1x_2}$ ) apuntaría una dirección de la causalidad entre las variables. La originalidad y sencillez de este diseño lógicamente ha posibilitado la creación de variedad de modelos apoyados sobre este patrón metodológico: diseños de paneles de dos variables en dos momentos (2W2V), diseño de dos variables en dos momentos con dos momentos con dos estimadores diferentes para cada variable (2W2V2I), diseños múltiples a base de varios paneles, distintos momentos de medición y diversos indicadores de las variables... (Kenny, D.A., 1979, Rogosa, D., 1979, Cox, T.H., y Campbell, D., 1979, Bentler, P., 1983).

Aunque muchos autores apoyan este tipo de modelos, en la práctica tienen ciertas limitaciones. De una parte, las relativas a la estabilidad de los coeficientes de correlación entre las variables en las mediciones sucesivas (erosión temporal) y la homogeneidad de las varianzas constituyen serios problemas a la afirmación de la dirección causal a partir de estos modelos. De otro lado, en ocasiones, las diferencias entre los coeficientes cruzados resultan significativas —a pesar de que apenas se diferencian— debido al tamaño de la muestra, pero no se tienen en cuenta los propios efectos del muestreo (sampling) y de la medición (testing). Además, al partir de una estadística univariada excluyen la utilización de variables latentes, por lo que cuando se utilizan varios indicadores por variable los análisis para obtención de los coeficientes de comparación se multiplican considerablemente (siendo  $K$  el número de momentos/- olas,  $p$  los indicadores de una variable ( $x$ ) y  $q$  el número de indicadores de la otra ( $y$ ), el total, de comparaciones sería igual a  $pq(2_k)$ .

La alternativa actual en el terreno de los métodos longitudinales correlacionales son los *modelos causales longitudinales*, a través de los cuales se pretende superar las limitaciones de los diseños cross-lagged y poder llegar a la explicación causal. En síntesis, se puede decir que estos diseños tratan de comprobar una serie de hipótesis teóricas que se formulan como posibles explicaciones que justifican la varianza entre los fenómenos. Para ello, previamente, es necesario

construir el modelo explicativo y especificarlo en términos matemáticos susceptibles de ser comprobados a través de un sistema de ecuaciones estructurales, cuya resolución ratificará o no la validez del modelo. Los diseños son muy semejantes a los de paneles de intervalos cruzados –se utilizan, incluso, los mismos códigos (2W2V, 2W2V2I, 3W2V3I,...)– sólo que parten de supuestos diferentes (no causación simultánea, medida de variables sin error y residuos no correlacionados). La plasticidad de estos modelos y la potencia de las técnicas de análisis –utilización de variables latentes, control de efectos indirectos, control de errores de medición...– permiten superar muchas de las limitaciones metodológicas anteriormente apuntadas y abren nuevos caminos a la investigación.

No obstante este tipo de diseños también tiene sus limitaciones. Entre las objeciones que habitualmente se formulan cabe destacar la no especificación en los modelos de variables importantes, el desajuste en la operativización y medición de las variables latentes, la fiabilidad/estabilidad de las medidas y sobre todo el problema de la estimación de los efectos recíprocos, para el cual aún no existe una clara solución desde la metodología correlacional (Rogosa, D., 1979, Pedhazur, E.Z. 1982).

A modo de resumen podríamos concluir que las opiniones de los investigadores en relación con la utilización de estos diseños no son unámines. En términos generales parecen decantarse a favor de los modelos basados en variables manifiestas –cross-lagged– como más apropiados para la descripción y predicción de los fenómenos, y por los modelos causales cuando se intentan su explicación dado que las variables latentes se ajustan más a los diseños teóricos a pesar de que en su estimación siempre hay un margen de error. En lo que sí coinciden la mayoría es que la utilización de estos modelos en las investigaciones longitudinales constituye en la actualidad la alternativa más esperanzadora (Crano, W., 1977, Jöreskog, K., y Sörbom, D., 1979, Kenny, D., A., 1979, Rogosa, D., 1979, Cox, T.H., y Campbell, D., 1979, Pedhazur, E.J., 1982).

## UN MODELO SECUENCIAL DE ANÁLISIS EN INVESTIGACIONES LONGITUDINALES.

La tendencia actual hacia la modelización causal en ciencias sociales está originando un evidente desfase entre el plano teórico –construcción de modelos, elección de variables, definición y medición de indicadores– y el desarrollo de las técnicas de análisis de datos. La oportunidad que tienen hoy los investigadores de utilizar sistemas de computación permiten efectuar análisis muy complicados que a veces adolecen no sólo de justificación teórica sino incluso de sentido común. En ocasiones se desprecian los análisis «clásicos» de datos cuando son imprescindibles para captar las distribuciones y relaciones entre las variables y poder formular un modelo teórico explicativo más ajustado. A la modelización causal debe llegarse después de una serie de análisis progresivos que le

permitan al investigador ir clarificando el campo, seleccionando las variables más representativas, determinando sus relaciones temporales y funcionales, etc..., es decir, conociendo de cerca la realidad que pretende explicar. De lo contrario de los modelos teóricos pueden quedar en simple artefactos que lo explican todo y no explican nada.

No existen modelos de referencia de cómo se debe proceder en estos análisis previos. Los autores preocupados por este tema (Jöreskog, K., y Sörbom, D., 1977, Nesselroade, J., y Baltes, P.B., 1979, Epstein, J., 1983, Baker, R., Mednick, B., y Broock, W., 1984) coinciden en una serie de puntos que no difieren sustancialmente del modelo secuencial que proponemos a título indicativo para el *análisis de datos en estudios longitudinales simples* con metodología correlacional y que hemos utilizado en un estudio de estas características que hemos realizado últimamente (De Miguel, M., y col. 1985). En síntesis los principales pasos de este modelo son los siguientes:

**1. Análisis de Distribuciones.** Una vez obtenidos los datos y organizadas las matrices, el primer tipo de trabajo a realizar es el análisis de las distribuciones lineales y cruzadas en función de las principales variables antecedentes e intermedias. A pesar de sus limitaciones (Kerlinger F. y Pedhazur, E.J., 1973), este tipo de técnicas permiten observar las anomalías en las distribuciones e interrelaciones de determinadas variables entre sí, establecer tablas de expectativas y/o probabilidad de ocurrencia de un suceso (predicción), estimar precozmente la epidemiología de deficiencia y/o alteraciones y sobre todo, establecer una ordenación de los análisis posteriores (Baker, R. Mednick, B., y Brock, W., 1984). Aunque estos análisis son diferentes según se trate de variables discretas o continuas, en la actualidad el desarrollo de las metodologías cualitativas ha dado gran impulso a estas técnicas: Markow, Wiggins, Coleman, Log-linear, etc. (Markus, G., 1979).

**2. Análisis Estructurales.** Dado el elevado número de datos que se manejan en los estudios longitudinales previamente a cualquier tipo de análisis es necesario proceder a una reducción/simplificación, y a estimar los cambios cualitativos que se producen en su estructura en las sucesivas estimaciones/evaluaciones. La técnica usual para ello es el análisis factorial que no sólo nos permite conocer las dimensiones subyacentes en la información obtenida en cada uno de los sondeos, sino que además, a través de modelos exploratorios y confirmatorios, posibilita seleccionar del conjunto de variables que entran en un análisis un número reducido representativo que pueda ser utilizado posteriormente en otros análisis con fines explicativos y agruparlas por bloques evitando aquellas que en las matrices manifiestan una gran multicolinealidad.

**3. Análisis sobre la estabilidad de coeficientes.** Otra de las estrategias de análisis de los estudios longitudinales es observar la estabilidad o inestabilidad en las relaciones/asociaciones bien entre un mismo tipo de variables y/o simila-

res en los distintos momentos de la exploración, bien entre determinadas variables que se consideran antecedentes en relación con otras (consecuente). A pesar de los fenómenos característicos de la erosión temporal, lo cierto es que el análisis de la estabilidad en la dirección y magnitud que presentan los coeficientes constituye un índice muy útil para captar las variaciones en la influencia que ejerce el entorno y las características de orden personal en los procesos de cambio individual.

**4. Análisis sobre las varianzas temporales.** Una vez identificados los sujetos en relación con la media muestral, los análisis se centran en estimar cómo afectarán los cambios en las variables más representativas sobre los resultados de los sujetos. Normalmente tales análisis parten de una estimación de la puntuación de los sujetos de una situación determinada respecto a otra futura, con la cual posteriormente se cotejan los resultados. De esta forma a través del análisis de regresión y de la varianza se pretende explicar las alteraciones entre los ajustes –puntuaciones esperadas y alcanzadas– en función de las variables diferenciales introducidas en cada análisis, especialmente aquellas procedentes del medio familiar y escolar (Rutter, M., y col. 1979, Madaus, G., y col. 1982). A pesar de las limitaciones y errores en la interpretación de resultados, últimamente, los avances de estas técnicas de análisis han sido considerables.

**5. Análisis sobre la secuencia temporal de las variables.** En el análisis de regresión la varianza asociada con las subpoblaciones se confunde con la varianza asociada con la muestra total por lo que se debe intentar obtener una separación entre ambas varianzas. Una de las formas de obtener tal distinción es observando si el poder discriminante de las variables predictoras es el mismo para cada una de las subpoblaciones. Para ello se establece una composición y ordenación temporal de las variables con mayor poder discriminante en cada uno de los momentos de estudio y sobre cada uno de los resultados estimados (variables output). La representación gráfica de tal composición –en forma de secuencia– pone de relieve las relaciones temporales entre las variables, y el posible papel que juegan en una relación causal.

**6. Análisis Causal.** Los análisis anteriores, aunque importantes, siguen quedando en el plano exploratorio. Por tanto si queremos comprobar posibles relaciones causa-efecto entre las variables debemos proceder, finalmente, a la utilización de diseños cross-lagged y modelos causales que nos permitan la afirmación o negación de la explicación causal. Ya hemos avanzado que desde el punto de vista metodológico son técnicamente más correctos los modelos causales dado que utilizan variables latentes, simplifican su operativización a través de múltiples indicadores, y analizan sus relaciones en términos de regresión, lo que permite aislar la influencia de otras variables. Pero, a su vez, los cross-lagged dada la sencillez de su diseño y su fácil resolución resultan muy apropiados para que los investigadores los utilicen inicialmente a efectos de descubrir

y/o descartar posibles relaciones causales, que posteriormente puedan ser reafirmadas a través de modelos más elaborados. En ese sentido se pueden concebir los análisis cross-lagged como un paso intermedio entre la simple correlación y los modelos estructurales.

Estos son los pasos o estrategias exploratorias de análisis de datos que entendemos deben guiar todo estudio longitudinal simple orientado a descubrir relaciones causales entre variables no controladas. En síntesis coincide con las etapas propuestas en otros modelos para demostrar la causalidad en ciencias sociales (Kenny, D., A. 1979), a saber: a) replicación de relaciones transversales b) descubrir relaciones en el tiempo entre posibles causas y efectos c) descubrir las diferencias cross-lagged y d) experimentar utilizando la variable causal como tratamiento. Este parece ser el camino a seguir en los análisis hacia la explicación causal.

## DIFICULTADES COMUNES DE LOS ESTUDIOS LONGITUDINALES.

Además de las dificultades conceptuales y prácticas inherentes a la realización de todo trabajo longitudinal, algunas de las cuales hemos avanzado (Harway, M., y Mednick, S., 1984), desde el punto de vista estricto de la metodología correlacional existen una serie de limitaciones asociadas a la utilización de estos modelos en el campo educativo que los investigadores deben conocer. A título indicativo señalamos algunas que se consideran fundamentales deducidas de los principales estudios longitudinales en el campo de la educación que anteriormente hemos citado:

**1. Dificultades derivadas de la selección muestral.** Dado que en este tipo de estudios se requiere la participación reiterada de unos mismos sujetos a lo largo del tiempo, existen serias dificultades para que la selección de los mismos se efectúe totalmente al azar. Aunque algunos autores sugieren técnicas específicas para paliar los efectos de la selección, lo cierto es que la mayoría de estos estudios parten de diseños pretests-protest tomando la maduración y/o experiencias de vida como única variable experimental/tratamiento, y utilizan muestras no representativas, por lo que quedan limitados en las técnicas de «control».

**2. Pérdidas selectivas de información.** Normalmente a lo largo de un estudio longitudinal existen pérdidas del conjunto de sujetos de la muestra y/o de la información que aportan los instrumentos de medida (attrition). El investigador debe observar si tales pérdidas se deben al azar o correlacionan con algún criterio de medida. Los resultados del estudio pueden ser engañosos si el investigador solamente utiliza en sus análisis los datos de los sujetos que completan toda la información a lo largo del ciclo temporal (Wohlwill, J., 1973, Labourie, E., y col. 1974), dado que raramente la mortandad experimental se produce totalmente al azar respecto a todas las variables del estudio.

**3. Efectos testing.** Una de las dificultades más comunes atribuibles a los estudios longitudinales es el efecto testing –aumento de puntuaciones en el post-test o modificación de conductas de los sujetos hacia la medida– producido por la aplicación reiterada de instrumentos de medida standarizados. Aunque los problemas conocidos como «techo y suelo» indudablemente introducen sesgos en tal medición, la dificultad más específica de los trabajos longitudinales es la equiparación de las medidas obtenidas por diferentes instrumentos que pretende evaluar una misma variable en dos momentos distintos. Cuando estas mediciones sucesivas se efectúan a corto plazo se producen los efectos testing, si se realizan a largo plazo aparecen los problemas de fiabilidad de los instrumentos (Campbell, D., y Stanley, J., 1973). El uso adicional de muestras de control dentro de cada cohorte y tiempo de medida es uno de los procedimientos utilizables para estimar este efecto.

**4. Erosión temporal.** Otra de las limitaciones de los estudios longitudinales es la selección de las variables apropiadas en función de los resultados que le interesan al investigador, precisamente porque en el diseño elaborado en los momentos iniciales no se preveen las relaciones/interacciones que establecen esas mismas variables con otras a lo largo del tiempo. Esto hace que determinadas relaciones/correlaciones entre variables se debilitan a medida que nos alejamos temporalmente –erosión temporal– lo que dificulta los planteamientos en términos explicativos. En ocasiones tal erosión tiene que ver con la aparición de una variable intermedia que el investigador debe aislar.

**5. Efectos de Agregación.** Existen además problemas derivados del tipo de unidad de análisis que se utiliza en el diseño. En muchos trabajos longitudinales se toma como tal unidad determinados colectivos –la escuela, la ciudad o el país– sin tener en cuenta que los individuos agrupados bajo tal unidad no están expuestos a las mismas condiciones educativas por lo que se produce un efecto de agregación de las varianzas. En ocasiones tal efecto queda aún más agravado y/o solapado por la «regresión estadística» que nos lleva a estimar las puntuaciones obtenidas en las mediciones sucesivas en términos de medias de población, hecho que se acentúa cuando nuestros instrumentos de medida presentan una baja fiabilidad.

**6. Diferencias relacionadas con las cohortes.** Los estudios longitudinales simples no captan los efectos derivados de la «historia», es decir acontecimientos del entorno específicos del período temporal en el que se realiza el trabajo, pero que no se encuentran en el siguiente. Lógicamente para aislar tales efectos el investigador debe introducir en el diseño la replicación utilizando la cohorte como variable a fin de conocer la dirección y proporción que aporta esta variable a la explicación del cambio temporal (Gleen, N., 1981). De ahí que se deba evitar la utilización de diseños simples a favor de estrategias secuenciales (Baltes, P.B., Reese, H., y Nesselroade, J., 1977, Baltes, P.B., Cornelius, S., y Nesselroade, J., 1978, Nesselroade, J., y Baltes, P.B., 1979).

## SUPERAR LA EVIDENCIA: UN RETO ABIERTO A LA INVESTIGACIÓN EDUCATIVA.

Durante los últimos años decir que las escuelas son ineficaces ha sido reiterar una evidencia sobradamente conocida desde la investigación y desde la realidad. Las teorías sociológicas sobre la escuela han resaltado esta tesis de tal modo que ha sido necesario que pasaran varios años para que los investigadores reexaminaran sus primeros informes y pudieran afirmar que las escuelas también eran diferentes. A partir de aquí comienza un nuevo camino en la investigación educativa.

El planteamiento actual es distinto. Para muchos de nosotros la escuela constituye una institución original cuyos efectos no pueden ser explicados por determinantes sociales y variables individuales. Probablemente nuestros métodos de análisis no sean adecuados o, al menos, en la actualidad no nos permiten asignar a cada variable un peso en la determinación del éxito. Pero esto no significa que la escuela sea una institución superflua. Al contrario, debe constituir el objetivo básico hacia el que deben orientarse los trabajos en el futuro, a fin de identificar las variables escolares intermedias entre origen social y éxito académico, tan importantes como desconocidas dado que en cada sujeto y realidad están tamizadas por factores específicos.

Parece ser que hoy las posibilidades que se abren con los modelos causales en investigaciones longitudinales permiten avanzar en este campo. Sin embargo nunca debemos olvidar que tales modelos deben interpretarse en el marco teórico en el que se formulan y que, por tanto, la obtención de un sistema de ecuaciones bien ajustadas algebraicamente no significa nada por sí mismo. Los resultados estadísticos obtenidos a partir de los modelos causales nos aportan las formulaciones exactas de las correlaciones y covariaciones entre las variables, pero no constituyen una prueba positiva de la verificación de las teorías. Su mayor utilidad está en que permiten establecer otras conexiones causales más coherentes y más próximas a la realidad. La investigación de las variables causales que determinan la efectividad de las escuelas no se resuelve con el diseño y comprobación del ajuste que se produce en las relaciones establecidas en un modelo teórico, sino a través de una búsqueda constante donde los resultados positivos de hoy nos ayudan a elaborar otros más ajustados a la realidad para mañana.

## REFERENCIAS BIBLIOGRÁFICAS

## A. BIBLIOGRAFÍA SOBRE LA EFICACIA DE LAS ESCUELAS

- AVERCH, H. y col. (1972). *How effective is schooling? A critical review and synthesis of research findings.* Rand Corporation, California.
- BIDWELL, C. (1975). *Nations, school districts and schools: Are there schooling effects anywhere?* in Meeting of American of American Educational Research Association. Washington DC.
- BLOOM, B. (1976). *Human characteristics and school learning.* McGraw Hill, New York.
- BRIDGE, R.; JUDD, C. y MOOCK, P. (1979). *The determinants of educational outcomes.* Mass Ballinger, Cambridge.
- BRIMER, A. y col. (1978). *Sources of differences in school achievement.* NFER Publishing, Slough.
- BROOKOVER, W. y col. (1978). «Elementary school social climate and school achievement.» *American Educational Research Journal*, 15.
- COLEMAN, J. y col. (1966). *Equality of educational opportunity.* Government Printing Office, Washington DC.
- COOLEY, W. (1974). Assessment of educational effects. *Educational Psychologist*, n.º 11.
- COMBER, L. y Keeves, J.; (1973). *Science education in nineteen countries. An empirical study.* Wiley, New York.
- CRAWFORD, J. y col. (1977). Classroom dyadic interaction: Factor structure of process variables and achievement correlates, *Journal of Educational Psychology*, 69.
- DE MIGUEL, M. (1984). «Investigaciones en torno a Educación Compensatoria.» *Revista Investigación Educativa*, 3.
- GOOD, T; BRIDDLE, J y BROPHY, J. (1975). *Teachers make a difference.* Holt, Rinehart y Winston, New York.
- HURN, C. (1978). *The limits and possibilities of schooling.* Allyn and Bacon, Boston.
- HUSEN, T. (ed). (1967). *International study of achievement in mathematics. A comparison of twelve countries.* Wiley, New York.
- JENCKS, C. y col. (1972). *Inequality: A reassessment of the effect of family and schooling in American.* Basic Books, New York.
- KELLACHAN, T. (1977). «Measuring school effectiveness», en SUMMER (ed.): *Monitoring national standards of attainment in schools.* NFER, Publishing, Slough.
- KLEIN, S. (1976). «Cost-quality research limitations. The problem of poor indices» en McDERMONT (ed). *Indeterminacy in education.* McCutchan, Berkeley, California.
- MADAUS, G. y col. (1979). *The sensitivity of measures of school effectiveness.* Harvard Educational Review, 49.

- MADAUS, G.; AIRASIAN, P. y KELLAGHAN, TH. (1980). *School effectiveness. A reassessment of the evidence*. McGraw Hill, New York.
- MARJORIBANKS, E. (ed.). (1974). *Enviroments for learning*. NEFER Publishing, Slough.
- MAYESKE, G y col. (1972 y 1973). *A study of our nations schools*. Departament of Health Education and Welfare, Washinton DC.
- MOSTELLER, F. y MOYNIHAN, D. (1972). *On equality of educational opportunity*. Vinage Books New York.
- PLOWDEN, B. (1966). *Children an their primary school*. H.M. Statimey Office, London.
- RAKOW, E. y col. (1978). «Assessing school and program effectiveness estimating teacher level effects.» *Journal of Educational Measurement*, 15.
- ROMBERG, TH.: (1981). «Methodology for the study schooling», in POPKEWITZ y TABACHICK (ed): *The study of schooling*. Praeger, New York.
- ROSENShINE, B. y FURST, N. (1971). «Research on teacher performance criteria» en SMITH (ed): *Research in teacher*. Prentice-Hall, Englewood Cliggs.
- SMITH, M.S. (1972). «Equality of educational opportunity», en MOSTELLER Y MOYNIHAN: *On equality of educational opportunity*. Vintage Books, New York.
- WALBERG, H. y RASHER, S. (1974). «Public school effectiveness and equality: New evidence and its implications.» *PHI Delta Kappan*, 56.
- WILEY, D. (1976). «Quany of schooling, a potent for policy» en SEWELL y col. (ed): *Schooling and achievement in american society*. Academic Press, New York.

## B. SOBRE ESTUDIOS LONGITUDINALES

- BAKER, R.; MEDNICK, B y BROCK, W. (1984). «An application of causal modeling techniques to prospective longitudinal data bases» en MEDNICK, HARWAY Y FINELLO (ed): *Handbook of longitudinal research*. Praeger, New York.
- BALTES, P. REESE, H y NESSELROADE, J. (1977). *Life-Span developmental psychology: Introduction to research methods*. Brooks-Cole, Monterey, California.
- BENTLER, P. (1983). *Structural equation models in longitudinal research*. University Press, California.
- COX, TH y CAMPBELL, D. (1979). *Quasi-Experimentation*. Rand McNally Publishing, Chicago.
- CAMPBELL, D. y STANLEY, J. (1973). *Diseños experimentales y quasiexperimentales en la investigación social*. Amorrortu, B. Aires.
- CRANO, W. (1977). «A cross-lagged panel analysis of selected data from the Berkeley growth study.» *Child Development*, 48.
- GLENN, N. (1981): *Cohort analysis*. Sage. London.

- HARWAY, M y MEDNICK, S. (1984). «Rationale for the study» en HARWAY, MEDNICK y FINELLO: *Handbook of longitudinal research*. Praeger, New York.
- HYMAN, H. (1955). *Survey design and Analysis*. The Free Press . New York.
- JORESKOG, K. (1978). «Structural analysis of covariance and correlations matrices». *Psychometrika*, 43.
- JORESKOG, K. y SORBOM, D. (1977). «Statistical models and methods for the analysis of longitudinal data», en AIGNER y GOLDBERGER (ed): *Latent variables in socioeconomic models*. North Holland, Amsterdam.
- JORESKOG, K. y SORBOM, D. (1979). *Advances in factor analysis an structural equation models*. Abt. Books, Cambridge.
- KENNY, D. A. (1979). *Correlation causality*. Wiley. New York.
- KERLINGER, F y PEDHAZUR, E. (1973). *Multiple regression in behavioral research*. Holt Rinehart-Winston. New York
- LABOUVIE, E y col. (1974). «On the internal and external validity of simple longitudinal desings». *Child Development*, 45.
- MARKUS, G. (1982). *Analysing panel data*. Sage. London.
- MEDNICK, S.; HARWAY, M y FINELLO, K. (1984). *Handbook of longitudinal research*. Praeger. New York.
- NESSELROADE, J. y BALTES; P. (eds). (1979). *Longitudinal research in the study of behavior and development*. Academic Press, New York.
- NESSELROADE, J. y REESE, H. (eds). (1973). *Life-Span development psychology methodological issues*. Academic Press, New York.
- PEDHAZUR, E.J. (1982). *Multiple regression in behavioral research: explanation and prediction*. Rinehart y Winston, New York.
- ROBINS, L. (1984). «Longitudinal methods in the study of development» en MEDNICK, HARWAY y FINELLO (ed): *Handbook of longitudinal research*. Praeger, New York.
- ROGOSA, D. (1979). «Causal models in longitudinal research» en NESSELROADE Y BALTES (ed): *Longitudinal research in the study of behavior and development*. Academic Press. New York.
- RUTTER, M. y col. (1979). *Fifteen thousand hours-secondary schools and their effects on schoolchildren*. Open Books, London.
- SCHAIE, K. y BALTES, P. (1975). «On sequential strategies in developmental research.» *Human Development*, 28.
- SCHAIE, K. y HERTZORG, C. (1982). «Longitudinal methods» en WOLMAN (ed.): *Handbook of development psychology*. Prentice-Hall, Englewood Cliffs.
- WALL, W y WILLIANS, H. (1970). *Longitudinal studies and the social sciences*. Heinemann, London.

WOHLWIEL, J. (1973). *The study of behavioral development*. Academic Press, New York.

**C. ALGUNAS INVESTIGACIONES LONGITUDINALES EN EL CAMPO EDUCATIVO.**

ASTIN, A. (1977). *Four critical years*. Jossey-Bass, S Francisco.

BAERT (eds). (1981). *Prospective longitudinal research*. University Press, Oxford.

BALTES, P.B. CORNELIUS, S.W. y NESSELROADE, J.R. (1978). «Cohort effects in Behavioral development: Theoretical and methodological perspectives.» En W. A. COLLINS (ed). *Minnesota symposium on child psychology* (v. II) N.J. LEA, Hillsdale.

BRONFENBRENNER, U. (1974). «A report on longitudinal evaluations of preschool programs.» *DHW Publication No. (PHD)* 75-25.

CONSORTIUM FOR LONGITUDINAL STUDIES.: 1977. *Lasting effects after preschool*. DHEW Publication No. (OHDS) 78.

DE MIGUEL, M. y col. (1985). *Rendimiento y Preescolarización. Estudio longitudinal a lo largo de la E.G.B. I.C.E.* Oviedo.

ENTWISLE, D. y HAYDUK, L. (1979). *Schooling and Young Children*. National Institute of Education. Londres.

ENTWISLE, D. y HAYDUK, L. (1982). *Early schooling*. J. Hopkins Press. Baltimore

EPSTEIN, J. (1983): «Longitudinal effects of person -family- school interactions on student outcomes» en KERCKHOFF (ed): *Research in Sociology of education and socialization*. DAI Press, Greenwich.

FLETCHER, J.; SATZ, P. y MORRIS, R. (1980). *The Florida longitudinal Project: A Review*. Alachua, Florida.

HAROWITZ, F. (1975). *Review of child development research*. University Press, Chicago.

JORDAN, T. (1979). *Prospective longitudinal studies of retarded children. Baby study*, St. Louis.

KRAUS, P. (1973). *A longitudinal of children from kindergarten into the adult years*. Wiley-Interscience., New York.

MEDNICK, S. y BAERT, A. (1981). *An empirical basis for primary prevention: prospective longitudinal research in Europa*. Oxford University Press. New York.

PACE, C. (1979). *Measuring outcomes of college*. Jossey Bass. San Francisco

PARKER, R. (ed). (1977). *The preschool in action: Exploring early childhood programs*. Allyn y Bacon, Boston.

RYAN, S. (ed). (1974). *A report on longitudinal evaluations of preschool programs*. DHEW Publication.

RUBIN, R. y BALOW, B. (1978). «Prevalence of teacher identified behavior problems: A longitudinal study. *Excepcional Children* n.º 45.

STEEL, L; WISE, L.y ABELES, R. (1979). «*Projet TALENT*». Jossey Bass, San Francisco

WEIKART, D. y col. (1970). *Longitudinal results of the Ypsilant Perry Preeschool Project*. High/S-cope Foundation, Ypsilanti.

WESTINGHOUSE CORPORATION / OHIO UNIVERSITY (1969). *The impact of Head start experience on children's cognitive and affective development*. Clearing house, Springfield.

## Hacia una conceptualización del producto educativo

por  
*Arturo de la Orden*

La determinación, identificación, análisis, categorización y crítica de los modos y características de la investigación y evaluación del rendimiento educativo implica, como paso previo, la respuesta a la pregunta ¿qué se pretende investigar y evaluar? o, si se prefiere, ¿qué es el rendimiento educativo?. La respuesta a esta cuestión supone identificar como se conciben los resultados o producto de la educación.

A primera vista podría parecer ocioso cuestionarse acerca de lo que entendemos por producto de la educación. El producto es el resultado de una acción o de un proceso. En nuestro caso, sería el resultado del proceso educativo, lo que sucede a los estudiantes como efecto de su experiencia educativa. Sin embargo, la crítica más benigna que puede hacerse de esta primera aproximación al concepto del producto educativo es resaltar, por un lado, su excesiva generalidad, que no permite avanzar mucho en la caracterización del producto educativo; y, por otro, su limitada perspectiva al centrarse en el individuo dejando al margen la consideración de la proyección social de los efectos de la educación.

En efecto, el producto educativo es lo que acontece al individuo como consecuencia del proceso de intervención sistemática que llamamos educación. Pero, ¿qué le acontece?. En términos aún muy generales, podemos decir que adquiere nuevas formas de conducta, modifica sus modelos de comportamiento, y las predisposiciones que los subyacen, como efecto del aprendizaje. Estas conductas se integran en la personalidad del sujeto cuya formación, de acuerdo a un modelo, constituye el fin de la educación. En consecuencia, puede identificarse el producto de la educación con la personalidad del sujeto concebida como modelo estructural y funcional más o menos estable de comportamiento individual. En esta línea se concibe el producto de la educación en el contexto del modelo funcional que ofrece Castillejo, J.L., (1981), quien presenta el «output» como resultante operativa, activa, de los aprendizajes y puestas en acción integrados del sujeto. Ahora bien, por un lado, la personalidad es la resultante comporta-

mental de un conjunto de estímulos que desbordando ampliamente la acción sistemática de la educación formal, constituye lo que se denomina el ambiente de un individuo; y, por otro, está en proceso permanente de autoformación, lo que supone, a su vez, una permanente integración de los productos educativos en el proceso general de educación. En otras palabras, los outputs se transforman en inputs para producir nuevos outputs en un complejo proceso de interacción. En esta perspectiva, resulta difícil delimitar lo que constituye el efecto específico de la educación formal así como los campos propios del proceso y del producto de la educación en términos absolutos. Esta reflexión nos conduce inexorablemente a descartar la conceptualización de un producto absoluto y total de la educación que se identificaría con el valor de la trayectoria completa de cada vida humana, y a centrar nuestra indagación sobre un producto concebido como relativo y parcial, especialmente desde el punto de vista de su investigación, evaluación y control.

Por otra parte, los sistemas educativos reflejan los valores de la sociedad, y, por consiguiente, los factores socioculturales condicionan la educación. Lo que los estudiantes aprenden, como lo aprenden y para qué lo apreden está íntimamente vinculado al contexto sociocultural y al momento histórico en que se produce. En última instancia, el producto de la educación, la personalidad del educando como «patrón» integrado de comportamientos sería inexplicable sin referencia a las coordenadas sociales y culturales en que se desenvuelve; pero, al mismo tiempo, estos patrones de comportamiento, efecto de la educación, configuran el perfil sociocultural de la comunidad que, de esta forma, se constituye en cierto grado, pero en sentido estricto, como producto de la educación. Al considerar, pues, el producto de la educación no podemos limitarlo al resultado directo en cada individuo sino que hemos de ampliarlo a sus efectos indirectos en la estructura y dinámica social. Si se me permite una analogía con el ámbito biológico, podría decirse que, del mismo modo que el mecanismo genético «produce» directamente al individuo e indirectamente determina, perpetúa y posibilita la evolución de las características de la especie, la educación –a la que podemos referirnos, al menos metafóricamente, como mecanismo genético de las culturas– «produce» directamente la personalidad de cada hombre e indirectamente contribuye decisivamente a configurar (perpetúa y posibilita al cambio) los caracteres socioculturales.

El producto de la educación se presenta, pues, como una realidad compleja y, en cierto grado, evasiva. Esta complejidad del producto educativo está, sin duda, en la base de las limitaciones manifiestas en su conceptualización. Del mismo modo que hemos aludido a la dimensión individual-social del producto educativo podemos considerarlo en la perspectiva de su distancia temporal a la finalización de procesos educativos específicos. Parece evidente que el producto inmediato no puede identificarse con el modelo de conducta integrada –personalidad– que constituye el fin de la educación. Se trata de productos –conductas– más vinculados a situaciones específicas y, por tanto, menos generalizables. El producto mediato, por su parte, se identificaría con los efectos

combinados e integrados de amplios y diversos procesos educativos, culminando en la personalidad individual, como resultante general de todos ellos, y en su influencia indirecta en la estructura y dinámica sociocultural.

Esta primera clasificación se relaciona con el viejo problema planteado por los constructores y usuarios de pruebas educativas acerca de la distinción entre tests de rendimiento y de aptitud. Para algunos rendimiento y aptitud son dos conceptos perfectamente separados y separables, mientras que para otros no son más que dos nombres para designar el mismo construto (Whipple, G.M. 1918 y Shrader, W.B. 1980). Kelly, T.L., (1927), y con él muchos autores hasta hoy, considera injustificada la distinción entre ambos conceptos. Frente a ésta, existe otra tradición teórica y práctica (Kaiser, H.F., 1974; Carroll, J.B., 1974; y tantos otros) que sostiene la distinción operacional de los constructos.

Sin negar la relación entre aptitud y rendimiento, un gran número de autores contemporáneos (Angoff, W.H., 1980; Dwyer, C.A., 1982; Mehrens, W.A., 1982) adoptan distintos criterios para distinguir los conceptos y especialmente los tests utilizados para su medida. El rendimiento suele identificarse con los resultados de la instrucción formal, es decir, con el aprendizaje que tiene lugar como consecuencia de un curso bien definido de instrucción; mientras la aptitud sería la consecuencia de diversos conjuntos de experiencias de aprendizaje no determinados, difusos y de larga duración. En otras palabras, el rendimiento para estos autores está más estrechamente vinculado que la aptitud al tiempo de instrucción, pero ambos no difieren esencialmente, ya que se identificarían con los extremos de un continuo (Cronbach, L.J., 1970). Los tests de rendimiento miden, en un extremo, estrictamente los resultados de la educación. Las puntuaciones de los tests de aptitudes, en el otro extremo, son relativamente independientes de programas específicos de instrucción. En general, un test de aptitud cuanto más orientado al contenido de la instrucción, más útil resultará para predecir el éxito escolar futuro de la misma materia: pero será menos adecuado para predecir el aprendizaje en general.

En síntesis, los tests de rendimiento se utilizan básicamente para medir los resultados o producto inmediato de la instrucción previa, mientras que los tests de aptitud intentan predecir el éxito de futuros aprendizajes, midiendo productos mediatos de instrucción no especificada. Por otra parte, se supone que el rendimiento mejora con relativa rapidez por efecto de la instrucción; mientras que el incremento de la aptitud es más lento, incluso sometiendo a los alumnos a programas específicos de instrucción. En todo caso, aunque esta diferenciación puede tener un valor práctico, no se apoya en una clara teoría que permita generar unas reglas para distinguir aptitud y rendimiento en muchas situaciones reales. Los tests de rendimiento y aptitud para el aprendizaje miden conductas fundamentalmente adquiridas a través de experiencias educativas, formalizadas o no. Parece evidente que lo que llamamos rendimiento –producto inmediato de la educación– depende funcionalmente de la aptitud para aprender; pero esta aptitud, a su vez, es, en gran parte, efecto de rendimientos educativos específicos acumulados a lo largo del tiempo e integrados en estructuras cognitivas

—producto mediato de la educación— cuya identificación y medida permite predecir futuros rendimientos en el aprendizaje escolar.

## TIPOS DE PRODUCTO EDUCATIVO

Cruzando las dos dimensiones señaladas (mediato-inmediato e individual-social) del producto educativo, se obtienen los cuatro tipos representados en la siguiente tabla:

		Producto	
		Individual	Social
Producto	Inmediato	Inmediato-individual A	Inmediato-social B
	Mediato	Mediato-individual C	Mediato-social D

Analizaremos brevemente cada uno de estos tipos.

*Producto individual-inmediato.* Se corresponde en líneas generales con los aprendizajes escolares específicos y es el que ha sido objeto de estudio más sistemático y persistente. Aún así, su conceptualización y exploración es muy limitada. Más allá de los contenidos de enseñanza, las destrezas intelectuales básicas y determinadas actitudes, apenas si se han hecho incursiones esporádicas y clasificaciones con criterios eclécticos y pragmáticos.

La clasificación prevalente hoy en pedagogía, agrupa los productos educativos en cognitivos y no cognitivos. Esta simple dicotomización del producto educativo se mantiene, en opinión de Weis, J. (1980), porque, pese a la existencia de otras perspectivas en la concepción de las características humanas —incluida la tradicional división en la intelectual, emocional, social, estética, física..., etc., recientemente retomada con algunas modificaciones por Foshay, A.W., (1975)— el énfasis que la escuela pone en rendimiento— académico y las habilidades cognitivas básicas, inclina a calificar cualquier otro tipo de producto que desborda los límites de lo normalmente medido como no-cognitivo o afectivo. La identificación de lo «no convencional» con «no cognitivo» puede haberse acentuado con la extraordinaria difusión de las taxonomías de objetivos educativos (Bloom, B.S., y otros, 1956; Krathwohl, D.R., y otros, 1964; Gagne, R.M., 1977). Pueden constatarse intentos de ampliar la conceptualización del producto inmediato individual a conductas relevantes más allá del conocimiento de las materias del currículum y las destrezas intelectuales básicas. Tal es el caso de Tuckman,

B.W., (1975), quien al formular la llamada «hipótesis de los productos diferenciados» pretende ampliar los estrechos límites de la tradicional «comprobación del rendimiento» sugiriendo que deben evaluarse al menos estas cinco categorías de productos:

- Conocimientos específicos y su comprensión
- Conocimientos generales y su comprensión
- Pensamiento y resolución de problemas
- Actitudes y valores
- Conductas relacionadas con el aprendizaje (iniciativa, atención, tiempo dedicado a aprender, etc.).

Esta lista de productos y otros muchos que pudiéramos aportar (Gagne, R.M., 1972; Glaser, R. 1972; Lohnes, P.R., 1973; Bennet, N., 1979), aunque carecen de una justificación teórica o conceptual rigurosa, como base de sus sistemas de categorías, tienen el mérito de desplazar el marco de referencia del producto educativo más allá de su tradicional fijación al rendimiento académico. También merece citarse en esta línea el trabajo de Messick, S. (1979) en el que desarrolla un sistema de doce categorías de productos educativos no cognitivos y sus correspondientes medidas.

Pese a estos intentos, el producto individual inmediato, y, en gran parte, también el mediato, sigue identificándose fundamentalmente con el conocimiento de los contenidos curriculares y ciertas destrezas cognitivas elementales no sólo por los educadores en su tarea evaluativa diaria y en las características de las pruebas determinales de curso y de nivel (De la Orden, A., 1981 y 1982) sino también por los investigadores que han llevado a cabo proyectos ambiciosos de evaluación del producto educativo general o los efectos de amplios programas. Tal es el caso del conocido estudio de Coleman, J.S., y otros (1966) sobre igualdad de oportunidades que tomó como unidad básica de producto educativo el aprendizaje verbal, si bien también se midieron otras variables como autoconcepto y motivación escolar. Algo similar puede afirmarse de los estudios de evaluación del más importante proyecto americano de educación compensatoria conocido como «Head Start» y cuyo objetivo fue equiparar educativamente a los alumnos procedentes de estratos familiares y sociales desfavorecidos con los más afortunados mediante programas de educación preescolar, aunque el número de categorías de producto no convencionales fue mucho mayor (Research Triangle Institute, 1972; Bridgeman, B., y Shipman, V.C., 1975). Otros estudios evaluativos con énfasis en los conocimientos y destrezas intelectuales básicos son los llevados a cabo por Haney, W., (1977), Kennedy, M.N., (1977), Stallings, J., (1975), Stebbins, L.B., y otros (1977) sobre los efectos del programa Follow-Through, de carácter compensatorio también, pero dentro del período de escolarización formal en el nivel primario. Finalmente, además del proyecto de evaluación de los Programas Renovados en los Ciclos Inicial y Medio de EGB, emprendidos por el Ministerio de Educación y Ciencia, y cuyos resulta-

dos no han sido publicados, es necesario mencionar dos de los mayores esfuerzos de evaluación masiva del producto educativo: uno, a nivel Educational Progress, comenzado en 1969 con extensas muestras de alumnos de 9, 13 y 17 años, es decir, primera etapa de enseñanza elemental y comienzo y final de la educación secundaria. El producto educativo se concretó a conocimientos y destrezas intelectuales en ciencias, matemáticas, estudios sociales, música, lectura, escritura, literatura y civismo (NAEP, 1975); otro, a nivel internacional, The International Study of Educational Achievement, llevado a cabo por la Asociación Internacional de Evaluación Educativa (IEA) con cientos de miles de alumnos de 22 países de Europa, América, Asia y Oceanía, con edades de 10, 14 y 16 años. El producto educativo medido fue conocimientos en siete materias, matemáticas, comprensión lectora, ciencias, literatura, geografía y francés e inglés como lenguas extranjeras (Härnqvist, K., 1975; Husen, T., 1967).

*Producto individual-mediato.* Como afirmaba más arriba, se trata de resultados educativos no vinculados a un proceso de intervención específico, que son consecuencia de integración de aprendizajes a distintos niveles de organización personal cuya resultante son aptitudes, predisposiciones, tendencias y orientaciones generales moduladoras de la conducta del individuo en campos y situaciones cada vez más amplios y que se manifiestan con progresiva independencia del contexto, culminando en la personalidad del individuo como estilo personal de percibir, de pensar, de actuar de relacionarse, en definitiva de ser, que se manifiesta en todos sus actos. Este producto podría identificarse con los niveles señalados por Castillejo, J.L., (1981) en su modelo funcional del proceso educativo como de alta estabilidad, es decir, utilizando su nomenclatura, actitudes, constelaciones de actitudes, actitudes radicales y personalidad «in fieri». Estilos cognitivos, tendencias estéticas, actitudes políticas, dimensiones de la personalidad, etc., son ejemplos de constructos que conceptualizan el producto individual mediato de la educación. La clasificación en cognitivos y no cognitivos carece de sentido. Como toda la conducta humana, tienen una base cognitiva y comportan una carga afectiva, involucrando a la persona en su totalidad.

Una forma inadecuada de conceptualizar el producto individual mediato de la educación es la implicada en las investigaciones sobre los efectos escolares tomando como variable dependiente la permanencia activa de los conocimientos y destrezas básicas adquiridos en la escuela una vez que ha transcurrido un período variable de tiempo (1, 5 ó 10 años). En realidad, aquí se confunde el producto mediato con el grado de retención del inmediato, cuando, en realidad, la virtualidad operativa de los aprendizajes y aptitudes específicas radica más en su contribución al desarrollo de aptitudes y predisposiciones conductuales generales que en su reproducción idéntica en situaciones posteriores. Cuando en un momento dado, una persona produce una respuesta específica, puede ciertamente explicarse como repetición, con ayuda del mecanismo mnemónico, de un aprendizaje singular anterior; pero, dada la multiplicidad de respuestas específicas no directamente aprendidas que el hombre produce permanentemente,

resulta más plausible considerarlas como manifestación concreta de modos generales de comportarse generados por la integración de aprendizajes directos en niveles de organización personal más altos merced a los mecanismos de transferencia.

Aunque existe una gran preocupación por el estudio de variables tales como orientación valorativa, creatividad, estilos cognitivos, actitudes, dimensiones de la personalidad, etc., expresiones típicas, en mi opinión, de productos individuales mediatos de la educación, en muy raras ocasiones se conceptualizan como tales, considerándolos más bien como rasgos personales sin una vinculación efectiva al proceso educativo. Cuando el diseño curricular e instructivo y el desarrollo de la acción educativa contemplen este tipo de constructos conductuales vinculados a los aprendizajes específicos y a la forma, secuencia y contexto en que éstos se producen, estaremos iniciando una vía fecunda en la sistematización y tecnificación de la intervención educativa en función de una jerarquía de objetivos que vaya más allá del aprendizaje de respuestas específicas.

*Producto social.* La educación aunque, a través del aprendizaje, se realiza en cada individuo, cumple, como antes afirmábamos, funciones sociales. Sociólogos, políticos y economistas se preocupan de la educación en la medida en que ésta influye en las estructuras, dinámica y desarrollo sociales. La influencia que la educación pueda tener en la movilidad social (Alexander, K., y Eckland, B.K., 1973), en el tipo y nivel de ocupación y en la distribución de la renta (Sewell, W.H., y Hauser, R.M., 1974; Tachibanaki, T., 1979), en la producción y productividad económicas a través de la formación de capital humano (Becker, G.S. 1975; McMahon, W.W., 1974; Taubman, P., 1976; Romero, J.L., y De Miguel, A., 1969; Vaizey, J., 1975), en la estratificación social y el logro de «status» (Treiman, D.J., 1977; Hout, M., y Garnier, M.A., 1979; Finney, J.D., 1972; Curry, E.W., y otros, 1978), en la reproducción ideológica (Gintis, H., 1972), y en tantas otras variables, se considera aquí como producto social de la educación, que, a su vez, condiciona el proceso educativo a través de un feedback permanente.

Por tratarse de un producto indirecto, consecuencias de la salida (output) del sistema educativo en su ambiente, ha sido más objeto de estudio y conceptualización en las ciencias sociológicas y económicas que en las pedagógicas. La consideración del producto social quizás pese más en los planificadores y administradores del sistema educativo que en los educadores mismos. En todo caso, el diseño de la intervención educativa no puede hacerse al margen de la consideración de las necesidades y valores sociales, entre otras razones porque se hallan implícitos en las demandas educativas de la sociedad a que la escuela necesariamente ha de responder.

## PROBLEMAS Y PERSPECTIVAS

Uno de los problemas capitales de la medida del producto educativo radica,

como se desprende de nuestra aproximación a su conceptualización, en la identificación y definición de las variables que lo representan adecuadamente. Generalmente, las variables utilizadas han sido elegidas respondiendo a criterios pragmáticos de carácter ecléctico. Parece aceptado unánimemente que es necesario evaluar la información incluida en el currículum básico, así como las aptitudes las aptitudes lingüísticas y matemáticas y otras destrezas intelectuales básicas. Esta unanimidad empieza a desvanecerse cuando se trata de precisar que aspectos de la información han de ser objeto de control. Parece que la cantidad o extensión de los contenidos dominados constituye una variable común aceptada y empleada; pero otras facetas como exactitud, dificultad o velocidad de respuesta, señalados por Brueckner, L.J., y Bond, G.L., (1961) son menos atendidas. Esta situación se constata no sólo en la evaluación realizada por los educadores (De la Orden, A., 1981 y 1982) sino también en los estudios de investigación. Glasman, N.S., y Biniaminov, I., (1981) en su revisión de los análisis input-output de las escuelas presentan 29 trabajos realizados básicamente en los últimos diez años y todos incluyen como variables de producto una, dos o las tres siguientes: lectura, aptitud verbal, aptitud matemática. Once de las investigaciones utilizan otra variable de tipo cognitivo que denominan puntuación compuesta de rendimiento y que pretende expresar el logro académico general en las materias del programa. Si volvemos la mirada a lo que se considera convencionalmente productos no cognitivos, sólo B trabajos incluyen una o varias de las siguientes variables: autoconcepto, aspiraciones educativas, «locus» de control, hábitos de estudio y mortalidad escolar (fracasos en general).

Parece claro que el énfasis evaluativo recae sobre aquellos aspectos que hemos caracterizado como producto individual inmediato, y por tanto el menos estable, y ello sin una teoría que respalde la categorización de dicho producto y la consiguiente elección y definición de variables. Esto supone que no existe garantía de que tal evaluación sea completa, aún dentro de su reducido ámbito. En cuanto a la evaluación del producto individual mediato, la situación resulta aún menos satisfactoria. La aplicación asistemática de pruebas de variables supuestamente no cognitivas es el único intento destacable. Aquí la falta de una teoría o modelo del proceso y del producto educativos es determinante, ya que al carecer del apoyo directo en las especificaciones de los programas, la elección de variables es aleatoria y arbitraria, dependiendo de la disponibilidad de instrumentos y de las preferencias del educador o evaluador.

Las modernas tendencias psicológicas ofrecen una interpretación de los procesos intelectuales que pueden ayudar a describir las metas y productos educativos en el lenguaje de la cognición, facilitando, como afirman Haertel, E.H., y Calfee, R.C., (1983), la formulación de los mismos de una forma tal que asegure la congruencia entre los contenidos de las disciplinas y los procesos mentales. En apoyo de esta tesis citan un conocido trabajo de Snow, R.E., (1980) para quien la psicología cognitiva ofrece una concepción del rendimiento educativo más rica y comprensiva que la psicología educativa tradicional. La perspectiva del aprendizaje complejo pone el énfasis hoy en la organización, representación

y uso del conocimiento a largo plazo, no en su acumulación inmediata. El producto de la educación no puede ya identificarse con la adquisición de datos destrezas específicas vinculados a un determinado contenido. Por el contrario, lo más significativo del aprendizaje es la permanente combinación, recombinación, estructuración y unificación del cuerpo de conocimientos que se van acumulando en sistemas funcionales para su utilización en los procesos de pensamiento y en aprendizajes posteriores. De este modo, el rendimiento educativo se presenta como una función de organización y no sólo como una función de adquisición. Se trata, pues, de identificar el producto de la educación con un conjunto de estructuras cognitivas y núcleos organizados de conocimiento que facilitan la solución de problemas y la comunicación dentro de las diversas materias de enseñanza.

Esta concepción del rendimiento educativo supone un cambio radical en su medida y, en consecuencia, en la investigación básica y aplicada en que sea considerado como variable relevante.

Otro problema de envergadura, relacionado con el anterior, es la dificultad de conocer la validez de las medidas del producto educativo. Algunos instrumentos de evaluación de productos inmediatos ofrecen índices de validez de contenido y de criterio (predictiva), aunque este no es el caso general. Pero el problema más grave es la dificultad de conocer la validez de constructo (Cronbach, L.J., y Meehl, P.R., 1955), especialmente relevante en la medida del producto individual mediato. En realidad, se trata de la validez de las variables o, mejor, de los constructos, que subyacen las variables representativas del producto educativo. Implica el hallazgo de evidencia que apoya la teoría que sirve de base a la variable tanto como a la medida misma. La validez de constructo expresa que las variables derivadas de una teoría del producto educativo son adecuadas. Una prueba tendrá validez de constructo en la medida en que sus resultados permitan inferencias sólidas sobre la variable medida cuyas características comportamentales, coherentes con la teoría, han sido empíricamente contrastadas. El problema, en nuestro caso, es que, en general, las categorías del producto educativo, es decir, las variables que los representan, no han sido derivadas de una teoría sino más bien seleccionadas de forma ecléctica y pragmática. En consecuencia, resulta difícil hoy hablar de validez de constructo de los instrumentos de medida del producto educativo; y, en el caso de las variables del producto individual mediato, esta situación se agrava, dado que las variables que lo representan son, por definición, constructos complejos de un alto nivel de abstracción. Haertel (1985) apunta en un reciente trabajo una línea de investigación sobre los resultados educativos considerados como constructos apoyados en teorías psicológicas que conciben los objetivos de la educación, más allá de su descripción conductual, en términos de procesos cognitivos o mecanismos de memoria que han de ser desarrollados a través de la instrucción (Greeno, J.G., 1980; Snow, R.E., 1980). En esta perspectiva, la validación de constructo, especialmente para la interpretación del producto de la educación en base a su medida mediante el uso de test de referencia criterial, es una necesidad inesquivable y

puede abrir caminos prometedores hacia una conceptualización más adecuada a los resultados de la instrucción.

La evaluación de los productos educativos de carácter social es, por ahora, más un deseo que una realidad. Los esfuerzos de sociólogos, economistas y administradores han sido en general desalentadores. El carácter indirecto de este producto y su manifestación a largo plazo, por un lado, y la dificultad de aislar en el contexto social los efectos de la educación de aquellos debidos a otros factores (económicos, políticos, culturales, etc.) plantea problemas educativos cuya solución no se vislumbra por ahora. Hay intentos alentadores en esta línea tendentes a identificar indicadores sociales de tales productos pero los resultados tanto teóricos como prácticos no parecen relevantes.

#### REFERENCIAS BIBLIOGRÁFICAS

- ALEXANDER, K. y ECKLAND, B.K. (1973). *Effects of education on the social mobility of high school sophomores fifteen years later* (1955-70). University of North Carolina. Chapel-Hill.
- ANGOFF, W.H. (1980). *Distinctions between aptitude and achievement tests*. En «Issues in testing». Educational Testing Service. Princeton, New Jersey.
- BECKER, G.S. (1975). *Human capital: A theoretical and empirical analysis white special reference to education*. (2nd edit.). Columbia Univsersty Press, New York.
- BENNET, N. (1979). *Estilos de enseñanza y progreso de los alumnos* Morata, Madrid.
- BLOOM, R.S. y otros (1956). *Taxonomy of educational objectives: The classification of educational goals*. Handbook I. Cognitive domain. David McKay, New York.
- BRIDGEMAN, B.y SHIPMAN, V.C. (1975) *Predicted value of measures of self-esteem and achievement motivation in four-tonine year old low-income children*. (ETS Technical Report P.R. 75-24). Educat. Testing Service. Princeton, New Jersey.
- BRUECKNER, L.J. y BOND, G.L. (1961). *Diagnóstico y tratamiento de las dificultades en el aprendizaje*. Rialp, Madrid.
- CARROLL, J.B. (1974). *Fitting a model of school learning to aptitude and achievement data over grade levels*. En D.R. Green (Ed.), «The aptitude-achievement distintion». CTB / McGraw-Hill, Monterey, Ca.
- CASTILLEJO, J.L. (1981). *Modelo funcional del proceso educativo*. En J.L. Castillejo y otros, «Teoría de la educación», Anaya/2, Madrid, págs. 99-137.
- COLEMAN, J.S. y otros (1966). *Equality of educational opportunity*. U.S. Government Printing Office. Washington, D.C.
- CRONBACH, L.J. (1970). *Essentials of psychological testing*. 3rd. edit. Harper and Row, New York.

- CRONBACH, L.J. y MEEHL, P.R. (1955). Construc validity in pyschological tests. *Psychological Bulletin*, Vol. 52, pp. 281-302.
- CURRY, E.W. y otros (1978). *Significant other influence of career decisions of black and white female urban youth*. Ohio State University Press, Columbus, Ohio.
- Dwyer, C.A. (1982). *Achievement testing*. En «Encyclopedia of Educational Research». 5th edit. AERA. MacMillan, New York.
- FINNEY, J.D. (1972). *Indirect effects of path analysis*. Sociological Methods and Research, Vol. 1, pp. 175-186.
- FOSHAY, A.W. (Ed.) (1975). *Essays on curriculum*. Teachers College Public, New York.
- GAGNE, R.M. (1972). *Domains of learning*. Interchange, Vol. 3. pp. 3-13.
- GAGNE, R.M. (1977). *The conditions of learning*. 3rd edit. Holt, New York.
- GINTIS, H. (1972). Towards a political economy of education: A radical critique of I. Illich's deeschooling society. *Harvard Educational Review*, Vo. 42, pp. 71-96.
- GLASER, R. (1972). Individuals and learning. The new aptitudes. *Educational Research*, Vol. 1, pp. 5-13.
- GLASMAN, N.S. y BINIAMINOV, I. (1981). Input-output analysis of schools. *Review of Educational Research*, Vol. 51, pp. 509-540.
- GREENO, J.G. (1980). Psychology of learning, 1960-1980: One participant's observation. *American Psychologist*, 35, 713-728.
- HAERTEL, E. (1985). Construct validity and criterion-refenced testing. *RER*, 55 (1), pp. 23-46.
- HAERTEL, E.H. y CALFEE, R.C. (1983). School achievement: Thinking about what to test. *J.E.M.* 20 (2), 110-132.
- HANEY, W. (1977). *The Follow-Through Experiment: Summary of an analysis of major evaluation reports*. Curriculum Inquire, Vol. 7, pp. 227-257.
- HARNQVIST, K. (1975). *The international study of educational achievement*. En N.L. Kerlinger (Ed.), *Review of Research in Education*, No. 3. AERA. F.E. Peacock, Itasca, Illinois.
- HOUT, M. y GARNIER, M.A. (1979). *Curriculum placement and educational stratification*. Trance Sociology of Education, Vol. 52, (3), pp. 146-155.
- HUSEN, T. (1967). *International study of achievement in mathematics: A comparison of twelve countries*. Almqvist and Wiksell. Stockholm.
- KAISER, H.F. (1974). *The chaldeans speak: An interpretive summary*. En D.R. Green (Ed.), «The aptitude-achievement distintion». CTB/ McGraw-Hill, Monterey, Ca.
- KELLY, T.L. (1927). *The interpretation of educational measurement*. World Book, New York.
- KENNEDY, M.N. (1977). *The Follow-Through Program*. Curriculum Inquire, Vol. 7, pp. 183-208.

- KRATHWOHL, D.R. y otros (1964). *The taxonomy of educational objectives: The classification of educational goals*. Handbook II. Affective domain. David McKay, New York.
- LOHNES, P.R. (1973). Evaluating the schooling of intelligence. *Educational Researcher*, Vol. 2, pp. 6-11.
- McMAHON, W.W. (1974). *Investment in higher education*. Lexington Books, Lexington, Mass.
- MESSICK, S. (1979). «Potential uses of noncognitive measurement in education. *Educational Psychology*, Vol. 71, pp. 281-292.
- MEHRENS, W.A. (1982). *Aptitude measurement*. En «*Encyclopedia of Educational Research*». 5th edit. AERA. MacMillan, New York.
- N.A.E.P. (1975). *Math fundamentals: Selected results from the First National Assessment of mathematics*. U.S. Printing Office. Washington, D.C.
- ORDEN, A. de la (1981). El control de la evaluación educativa. Un ejemplo. *Boletín de la Inspección de Educación Básica del Estado*. N.º monográfico 1/2, Diciembre, págs. 63-99.
- ORDEN, A. de la (1982). Análisis de las pruebas terminales de EGB. *Revista Española de Pedagogía*, N.º 156, págs. 7-40.
- Research Triangle Institute (1972). *A report on two national samples of Read Star classes. Some aspects of child development of participants in full year 1967-68 y 1968-69 Programs*. R.T.I. Park. N.C.
- ROMERO, J.L. y DE MIGUEL, A. (1969). *EL capital humano. Ideas para una planificación social de la enseñanza en España*. Confederación Española de Cajas de Ahorros, Madrid.
- SCHRADER, W.B. (Ed.) (1980). *Measuring achievement: Progress over a decade. New directions for testing and measurement*. Jossey Bass, S. Francisco.
- SEWELL, W.H. y HAUSER, R.M. (1974). *Education, occupation and learnings: Achievement in the early career*. Department of Sociology, University of Wisconsin, Madison.
- SNOW, R.E. (1980). *Aptitude and achievement*. En W.B. Shrader (Ed.), «*Measuring achievement: Progress over a decade. New directions for testing and measurement*», Jossey-Basss, S. Francisco.
- STALLINGS, J. (1975). *Implementation and child effects of teaching practices in Follow-Through classrooms*. Monographs of The Society for Research in Child Development (Series, No. 163), Vol. 4, pp. 7-8.
- STEBBINS, L.B. y otros (1977). *Education and experimentation: A planned variation model*. Vol. IV-a, An evaluation of Follow Through. Abt. Associates. Cambridge, Mass.DS
- TACHIBANAKI, T. 1979). Models for educational and occupational achievement over time. *Sociology of Education*, Vol. 52 (3), pp. 156-162.
- TAUBMAN, P. (1976). «Earnings, education, genetics and environment. *The Journal of Human Resources*, Vol. 11, pp. 447-461.
- TREIMAN, D.J. (1977). *Occupational prestige in comparative perspective*. Academic Press, New York.

- TUCKMAN, B.W. (1975). The differentiated outcome hypothesis. *CEDR Quarterly*, Vol. 9, pp. 8-11.
- VAIZEY, J. (1975). Problemas de la economía de la educación. *Revista de Educación*, Vol. XXIII, N.º 326-237, págs. 15-22.
- WEISS, J. (1980). *Assessing nonconventional outcome of schooling*. En D. Berliner (Ed.), *Review of Research in Education*, No. 8. AERA, Washington, D.C.
- WHIPPLE, G.M. (Ed.) (1918). *The measurement of educational products*. The 17th Yearbook of the N.S.E.E. Part. II. Public School Publishing, Bloomington, Illinois.

# Modelos de investigación sobre el rendimiento académico. Problemática y tendencias

por  
*Sebastián Rodríguez Espinar*

## INTRODUCCIÓN

Creo que puede afirmarse que el tema del rendimiento académico emerge en el panorama de la investigación educativa de nuestro país con algunos años de retraso por lo que respecta a un planteamiento con cierta seriedad metodológica, debiéndose señalar que aún seguimos careciendo de informes o estudios que vayan más allá de la mera contabilización de frecuencias de los "productos de certificación académica" (G.E., B.U.P., F.P., Licenciaturas, etc). Desde esta perspectiva, a nivel nacional, no debe sorprendernos la afluencia de "opiniones explicativas" que en los últimos tiempos se nos vienen ofreciendo acerca del fenómeno del fracaso escolar y de las alternativas al mismo, explicitadas en los proyectos de reforma propuestos.

Es cierto que estamos asistiendo a un movimiento generalizado de cuestionamiento de los planteamientos de los diferentes sistemas educativos que no han visto logradas las metas propuestas en las reformas introducidas en la década de los sesenta (1) a las que nosotros llegamos un poco tarde (aún no han terminado su ciclo universitario los alumnos que en Septiembre de 1970 comenzaron 1.º de E.G.B.). La polémica generada sobre los deseables "productos" de la educación es indicador claro de la presión de la sociedad sobre la institución educativa y que queda reflejada en las decisiones políticas de cambio.

La compleja polémica sobre el tema del rendimiento académico (éxito o fracaso escolar) viene determinada por una triple convergencia de dimensiones de la institución escolar que generan una serie de objetivos de difícil logro y que dan paso a consideraciones negativas acerca de los efectos de la propia institución educativa (Coleman *et al.*, 1966; Bernstein, 1970; Jencks *et al.*, 1972) (2). La primera dimensión aludida es la social. La acción de la escuela debe facilitar la nivelación de las desigualdades sociales. En este sentido, la productividad so-

cial de la misma vendrá determinada por la eliminación de diferencias en los logros vitales del individuo. Si un logro vital es la mejora del "status social", cuyo indicador básico es el salario, habremos de convenir en plantear el rendimiento en relación a la igualdad de acceso a las funciones productivas de la sociedad y a la recepción de los beneficios de tal producción (3).

La segunda dimensión es la educativa-institucional y se centra en dar respuesta a la pregunta: ¿En qué medida son adecuados los diferentes tratamientos educativos (programas, métodos, organización, profesorado, etc.) para el logro de los objetivos propuestos?. La minimización de las diferencias en la calidad e intensidad del rendimiento ha sido y es un reto constante aunque, como señalan Rutter *et al.* (1980), las diferencias de logro dentro de cualquier sistema escolar son evidentemente mayores que las diferencias entre sistemas (escuelas). Elevar la calidad de la educación no tendrá, y no debería tener, el efecto de hacer a todos los alumnos iguales. La razón es que una de las principales formas en que operan las diferencias individuales es en la habilidad de sacar provecho a las oportunidades ofrecidas, y esto tanto si las diferencias individuales reflejan variaciones genéticas como si son el resultado del ambiente.

La tercera dimensión es la económica. Esta hace referencia a la medida en que las inversiones en educación (su forma y contenido) producen la adecuada satisfacción a las demandas de la sociedad que genera los recursos que se aplican. Si queremos ser oídos en los círculos de toma de decisiones políticas no podemos olvidar los efectos económicos que conlleva cualquier planteamiento de reforma de la educación. En momentos de crisis económica, el concepto de eficiencia adquiere una significación especial como lo ponen de manifiesto los planteamientos de las investigaciones sobre el rendimiento englobadas bajo el epígrafe *FPE* (Función Productiva de la Educación) a las que haremos referencia más adelante.

En consecuencia, la investigación sobre el rendimiento académico nos plantea dos retos importantes. Por una parte, la previa determinación de los *productos* a considerar (es el reto de carácter teórico-filosófico sobre los objetivos de la educación). En segundo lugar, un doble reto de carácter metodológico: cómo evaluamos la calidad de dichos productos y cómo analizamos los factores determinantes de la calidad del rendimiento; es decir los modelos y metodologías de análisis de dichos determinantes.

Es obvio señalar que nuestro cometido en este III Seminario habrá de centrarse en la reflexión sobre las posibilidades de dar respuesta al doble reto metodológico planteado. De una parte, abordar la ya clásica y doble antinomia evaluación cualitativa vs. evaluación cuantitativa y evaluación criterial vs. evaluación normativa. En segundo lugar, debatir los diferentes modelos y metodología de análisis de los factores de rendimiento académico: sus posibilidades, adecuaciones y limitaciones. Todo ello a la luz de las numerosas aportaciones que se harán desde la vivencia concreta de la investigación realizada –única vía de ir más allá del nivel de "opinión" para lograr un lenguaje de comunicación que trascienda nuestras "convicciones" y se asiente en nuestro "saber"–.

Ahora bien, es difícil adentrarse en el análisis de los modelos de investigación sobre el rendimiento académico sin plantear, aunque sea de manera muy somera, el tipo de productos analizados en dicha investigación, ya que difícilmente podremos entender los aspectos metodológicos sin un adecuado análisis del criterio utilizado.

## EL CRITERIO DE RENDIMIENTO ACADÉMICO

La multidimensionalidad del concepto de rendimiento académico es indiscutible en el contexto de la investigación sobre el tema dada la pluralidad de efectos o logros (*outcomes*) perseguidos por toda acción educativa, pero, como señala Anderson (1982), pocos investigadores van más allá de considerar un par de dimensiones del mismo. Es obvio afirmar el valor indicativo de las dimensiones elegidas en relación a la preponderancia de determinados objetivos en la ejecución de un programa educativo. Ahora bien, no hemos de olvidar la presión social e institucional (certificación académica) a la hora de elegir el investigador unas u otras dimensiones.

Los trabajos de Schneider *et al.*, Sorensen y Hallinan, Mc Partland *et al.* y Levin (vid. Anderson, 1982, p. 383) ponen de manifiesto una serie de consideraciones:

- El uso de una sola medida de rendimiento (tests de rendimiento o calificaciones) en un momento determinado (final de curso, etapa o escolaridad) está infraestimando los efectos educativos de la escuela.
- Debería tenerse presente no sólo el aprendizaje inmediato, sino también el aprendizaje relevante al futuro del alumno (tanto educativo como ocupacional o vital).
- En los resultados de la enseñanza deben incluirse tanto los logros alcanzados en los objetivos académicos como en los de formación y desarrollo de la personalidad.
- Las actitudes ante el fenómeno de aprender, así como las conductas sociales de clara incidencia en la convivencia de la comunidad deberían tener un mayor peso en la determinación de la calidad del producto educativo.

Idénticas consideraciones nos ofrecen Centra y Potter (1980) al señalar que con excesiva frecuencia el rendimiento académico se ha centrado en las habilidades básicas como variables dependientes en la investigación, descuidando otras habilidades cognitivas de rango superior así como rendimientos no-cognitivos de la naturaleza del desarrollo moral, autoafirmación o el mismo éxito vocacional y ocupacional. La consideración de este tipo de criterio de rendimiento tal vez nos llevaría a identificar condiciones y estilos de enseñanza bastante diferentes a los investigados hasta el momento.

La amplia revisión efectuada por Glasman y Biniaminov (1981) sobre el modelo *input-output* puso de manifiesto que las 3/5 partes de las investigaciones usaron como criterio de rendimiento académico algún tipo de medida cognitiva

(bien test de rendimiento u otro tipo de test aptitudinal). La dimensión no-cognitiva abarcó aspectos como los de actitudes de los alumnos, aspiraciones educativas, tasas de asistencia y abandono, completar ciclos de enseñanza y paso al siguiente nivel, etc. Este último criterio, de paso de un ciclo a otro de enseñanza, ha recibido especial atención por los enfoques productivos de la educación (FPE) al ser considerado como claro indicador del efecto "igualador" de la educación (Schiefelbein y Farrell, 1982) y ha sido adoptado en las investigaciones que sobre el tema se han llevado a cabo en los países en desarrollo. No ha de extrañarnos, por tanto, que este sea el criterio "oficial" que actualmente se está barajando en nuestro país.

La crítica al excesivo énfasis en los objetivos "academistas" es clara. Gray (1981) señala que asumir que sólo hay una dimensión del rendimiento o efectividad de la escuela está totalmente injustificado, pues se ha extralimitado la creencia de que un alto rendimiento académico en la dimensión cognitiva lleva un alto rendimiento en otras dimensiones y, en consecuencia, la escuela que "produce" altos rendimientos de certificación académica es la que mejor "educa" en el sentido amplio y mayor del término educar. En la dirección de crítica al uso de pruebas de rendimiento referidas a la norma para otorgar dichas certificaciones se pronuncian Bridge *et al.* (1979, p. 47) al poner de manifiesto que las limitaciones de representatividad de sujetos con determinados tipos de "handicaps" limitan la validez de los hallazgos en cuanto a los factores determinantes del rendimiento de tal tipo de sujetos. Por otro tipo de razones, Madaus *et al.* (1979) se pronuncian en contra de la consideración de los "exámenes oficiales" (control institucional superior) como indicadores válidos del rendimiento (4). La investigación de Rutter *et al.* (1979) al considerar la tasa de delincuencia como un indicador de la efectividad de la institución escolar dió un toque de atención al "intelectualismo" de las investigaciones en Gran Bretaña, provocando algunas reacciones críticas acerca de la adecuación de tales criterios como medidas del rendimiento escolar (Acton, 1980).

En resumen, puede decirse que la delimitación de un adecuado criterio de rendimiento debe tomar en consideración dos tipos de componentes: a) un componente de contenido basado en los resultados tanto cognitivos como no-cognitivos, b) un componente temporal que abarque tanto el rendimiento inmediato (en el lapsus temporal de la permanencia del alumno en la institución educativa) como el rendimiento mediato que considere los logros vocacionales-ocupacionales así como los logros de tipo social (5). En el siguiente cuadro, y tomando como base el planteamiento de Astin (1984), ofrecemos una síntesis no exhaustiva de los diferentes componentes de la aceptada multidimensionalidad del concepto de rendimiento.

Es obvio suponer la dificultad de trasladar a cualquier diseño de investigación una adecuada operativización y medida de cada uno de los múltiples componentes de un adecuado criterio de rendimiento. De aquí, la enorme dificultad de proceder a una clasificación exhaustiva y excluyente de los modelos de análisis de los determinantes del rendimiento que han venido utilizándose en la in-

*Multidimensionalidad de los resultados de la enseñanza (RDTO)*

<u>Dimensión</u>	<u>Tipo de resultados</u>	
	<u>Cognitivos</u>	<u>No-cognitivos</u>
Psicológica	<ul style="list-style-type: none"> <li>-Habilidades básicas de aprendizaje (cómo aprender)</li> <li>-Aptitudes específicas</li> <li>-Razonamiento</li> <li>-Pensamiento crítico</li> <li>.....</li> <li>-Madurez vocacional</li> <li>-Expectativas de logros educativos y culturales</li> <li>-Nivel de responsabilidad en el trabajo</li> <li>-Status ocupacional y nivel de ingresos</li> <li>.....</li> <li>-Saberres culturales</li> <li>-Saberres instrumentales</li> <li>.....</li> </ul>	<ul style="list-style-type: none"> <li>-Autoconcepto</li> <li>-Actitudes y valores</li> <li>-Motivación para el rendimiento</li> <li>-Locus de control</li> <li>-Satisfacción práctica escolar</li> <li>.....</li> <li>-Habitos personales</li> <li>-Relaciones personales</li> <li>-Civismo y ciudadanía</li> <li>-Trabajo de/en grupo</li> <li>-Actitudes familiares ante el hecho e institución educativa</li> <li>.....</li> <li>-Expresión de sentimientos</li> <li>-Comprensión y conocimiento de sí mismo y de los demás</li> <li>.....</li> </ul>
Sociológica		
Instructiva		

vestigación educativa. Con esta premisa nos adentraremos en el tema de los modelos en la investigación sobre el rendimiento académico.

## MODELOS DE ANÁLISIS DE LOS DETERMINANTES DEL RENDIMIENTO ACADÉMICO

En un anterior trabajo (6) señalábamos la dificultad de explicitar una correcta clasificación de los modelos utilizados en la investigación sobre el tema del rendimiento dadas las interconexiones existentes entre los diferentes criterios de clasificación. Pese a ello (asumiendo el riesgo de posibles inadecuaciones), y aceptando la afirmación de Anderson (1982, p. 384) de que los diferentes modelos requerirán no solamente diferentes variables (en virtud de la teoría adoptada), sino también diferentes metodologías de análisis estadístico que pueden llevarnos a diferentes conclusiones, proponemos los siguientes criterios de clasificación:

### **Atendiendo a la naturaleza de los factores de rendimiento.**

Este ha sido el criterio más general utilizado ya que ha centrado su atención en el tipo de variables presumiblemente asociadas al éxito o fracaso escolar.

Tres grandes tipos podemos distinguir:

– De carácter psicológico: hacen referencia a los atributos personales inherentes tanto al alumno como al profesor. Características aptitudinales, de personalidad, intereses, etc. componen un amplio mapa tratado con exhaustividad en la literatura sobre el tema. La ya clásica obra de Lavin (1965), los trabajos de Cattell y Butcher (1968), Wasna (1974), Urouglu (1979), Mateo (1980), Pelechano (1977) y Rodríguez (1982) son, entre otros muchos, exponentes de la toma en consideración de este tipo de factores en su relación con el rendimiento escolar.

– De carácter sociológico: hacen referencia tanto a los elementos estructurales y estáticos del contexto (familiar, institucional y social) como a sus aspectos dinámicos, sobre todo al considerar a la escuela como un sistema social de relaciones entre la familia, profesores y alumnos. Las corrientes ambientalistas con sus diversas matizaciones ocupan un lugar importante en la investigación sobre el tema. A los ya clásicos informes de Coleman *et. al.* (1966), Jencks *et. at.* (1972) y Mayeske *et. al.* (1972) hemos de añadir los también clásicos planteamientos de Bloom (1964) y las aportaciones de la escuela británica representadas por las obras de Marjoribanks (1974, 1979) y Keeves (1972).

Hemos de reconocer la constante evolución de los planteamientos sociológicos en el tema del rendimiento académico, habiéndose pasado de una consideración estática a los mismos (clase social, nivel de estudios de los padres, zona de ubicación del centro, etc.) a poner la atención en los factores creados como consecuencia de las interrelaciones creadas entre el medio y los protagonistas del proceso educativo (enfoque ecológico). Merecen destacarse las aportaciones de Brookover *et. al.* (1979), Walberg (1979), Kurpius (1978) y Moos (1979). Una síntesis de estos enfoques puede encontrarse en el bien documentado trabajo de C.A. Anderson (1982) "The Search for School Climate: A Review of the Research".

– De carácter didáctico: hacen referencia a las características de actuación del docente como a la conducta puesta en juego por el discente, así como a la multitud de factores asociados a las estrategias, métodos, recursos, organización, etc. del proceso de aprendizaje. Hemos de reconocer la posición difícil en que se encuentran muchos de los planteamientos didácticos, en orden a demostrar su efectividad, cuando se "aprieta" la pinza explicativa ofrecida por los factores psicológicos y sociológicos en cuanto al éxito o fracaso escolar. De aquí la importancia del escrupuloso análisis metodológico al que hay que someter a muchos de los estudios explicativos de la "ineficacia" de la escuela. Antes de arrojar la toalla de la utilización de metodologías semejantes a las empleadas por los enfoques psicológico y sociológico se ha de hacer un notable esfuerzo para definir adecuadas variables didácticas y enmarcadas en un modelo explicativo al que se le puedan explicar las técnicas de análisis actualizadas. Algunas de las conclusiones derivadas de los trabajos presentados en la obra de Shulman (1977), así como las revisiones efectuadas por Hoge y Luce (1979), Centra y Polter (1980) y Doyle (1983) apuntan a la necesidad de los modelos causales y longitudinales como la vía claramente alternativa a la inconsistencia de muchos de

los hallazgos obtenidos con este tipo de factores. Ya tendremos ocasión de volver sobre este punto que considero nuclear en las discusiones de este III Seminario.

### **Atendiendo a la tipología del diseño de investigación**

En primer lugar nos encontramos con la tradicional dicotomía de estudios descriptivos (*ex-post-facto*) y estudios experimentales. No es el momento ni el objetivo abordar técnicamente las características de la diferenciación ni las vías de aproximación para encontrar un punto de encuentro en la explicación del fenómeno del rendimiento académico. El examen de la literatura sobre la investigación del rendimiento académico pone en evidencia la preponderancia de los estudios descriptivos, y ello por múltiples razones de todos conocidas: dificultades técnicas, coste, implicaciones éticas que conlleva la diferenciación de tratamiento en determinado tipo de variables, etc., etc. que están presentes en todo estudio experimental. Pese a muchos resultados descorazonadores en estudios experimentales (7), autores como Dyer (1972), Brookover *et. al.* (1979), Rutter *et. al.* (1979) insisten en la necesidad de plantear investigaciones experimentales a fin de comprobar los factores que alteran los resultados académicos de los alumnos.

En segundo término se plantean los estudios transversales (cross-sectional) frente a los estudios longitudinales. Las limitaciones de los primeros son evidentes, sobre todo para poder detectar el lento proceso acumulativo de los efectos de la acción educativa. Glasman y Biniaminov (1981, p. 535) afirman que las teorías del rendimiento académico seguirán siendo inadecuadas en la medida en que no se avance en el conocimiento de las interacciones entre factores así como de sus efectos longitudinales. En la última parte de nuestra ponencia plantearemos algunos de los problemas de análisis de los estudios longitudinales al presentar la alternativa de los modelos de ecuaciones estructurales al método de análisis de “*cross-lagged correlation*”.

### **Atendiendo a la función atribuida a la institución escolar**

Dos modelos, ya familiares, pueden ser considerados: de una parte el modelo de Entrada-Salida (*Input-Output*) que considera a la escuela como un sistema de atributos en donde es necesario analizar la relación existente entre determinadas entradas y sus respectivas salidas. El objetivo estriba en identificar las diferentes fuentes de diferenciación independientemente de los procesos relacionales generados en el propio proceso de aprendizaje. Este modelo, ligado al modelo aditivo que abordaremos más adelante, al utilizar fundamentalmente las técnicas de análisis basadas en la regresión, busca determinar las unidades de entrada que producen alteraciones en el producto. En consecuencia, consideraciones técnicas como el tema de las unidades de análisis así como los enfoques en el

tema de la partición de la varianza son aspectos que han de tenerse en cuenta para la correcta interpretación.

El segundo modelo es el denominado Proceso-Producto que considera a la institución escolar como «algo más que una caja vacía» y centra su atención en las nuevas variables generadas como consecuencia de la interacción de los comportamientos docentes y discentes. Por tanto, el modelo Proceso-Producto está en conexión directa con los modelos mediacionales e interactivos, aunque lamentablemente se han venido utilizando técnicas de análisis estadístico inadecuadas al planteamiento teórico subyacente. Remitimos a un anterior trabajo (vid. nota (6)) para la consulta de algunos de los hallazgos sobre el tema utilizando ambos modelos.

### **Atendiendo a la explicitación de relaciones entre variables**

Estrictamente hablando éste debe ser el criterio de determinación de modelos en el análisis de los determinantes del rendimiento. Acorde con la definición ofrecida por Snow (1973) de que un modelo es «un posible sistema de relaciones entre fenómenos expresadas en términos verbales, materiales, gráficos o simbólicos» (p. 81), la explitación apriorística de estas relaciones es consustancial con la adopción de una teoría explicativa del fenómeno del rendimiento académico, y está en la base de los análisis causales. Desde esta óptica, Anderson (1982) señala tres tipos de modelos: aditivos, mediacionales e interactivos, aunque desde mi punto de vista siguen existiendo interconexiones entre los dos últimos basadas en la adopción de una interacción o influencia unidireccional o bidireccional.

1.- *Modelos aditivos.* Asumen que las variables o factores, sean de la naturaleza que sean, influyen directamente en los resultados de los alumnos de manera independiente pero sus efectos son sumativos; es decir, el efecto final es la suma de los efectos parciales aportados por cada una de las variables tomadas en consideración. Como hemos señalado con anterioridad, las técnicas de análisis empleadas se basan en la regresión múltiple y sus limitaciones interpretativas han sido puestas de manifiesto por Luecke y McGinn (1975), así como en la comunicación presentada en el anterior Seminario (J. Mateo y S. Rodríguez, 1984).

Sin embargo, hemos de reconocer la influencia ejercida por este modelo –dada su utilización por los economistas de la educación– en el amplio panorama de investigaciones llevadas a cabo bajo los auspicios del Centro Internacional de Investigaciones para el Desarrollo con sede en Canadá y financiado por el Bco. Mundial. Es el modelo adoptado por la ya mencionada «Función productiva de la educación» (FPE), cuyo objetivo estriba en determinar las unidades necesarias de incremento en los factores de entrada a fin de que se produzca el incremento de una unidad en el rendimiento final. (8). Dada la repercusión de tales estudios en las esferas gubernamentales, se ha de poner especial atención a la extrapolación de conclusiones e inadecuación metodológica de muchas de las inferencias causales efectuadas.

Desde otra perspectiva de utilización de los modelos aditivos, éstos han prestado y prestarán una gran contribución como fase exploratoria para la formulación de modelos mediacionales e interactivos. Tal ha sido el proceso seguido por Glasman y Biniaminov (1981) y Centra y Potter (1980) para la propuesta de sendos modelos explicativos del rendimiento y que recogemos en el siguiente apartado. En esta línea, la utilización de las técnicas de síntesis cuantitativas (meta-análisis) ofrece una nueva perspectiva exploratoria, tanto por lo que se refiere a estudios correlacionales como a experimentales. A título informativo ofrecemos algunas referencias bibliográficas sobre trabajos de meta-análisis que pueden ser de utilidad en futuras investigaciones sobre el tema de rendimiento en nuestro país (9).

En términos generales señala Walberg (1974), los modelos aditivos han considerado el aprendizaje (rendimiento) como una función de las aptitudes, del ambiente y de las estrategias de instrucción;  $R = f(\text{Aptitudes} + \text{Ambiente} + \text{Instrucción})$  considerando a cada uno de estos factores como sumandos independientes. Ahora bien, no es menos cierto que cada uno de los «sumandos» ha recibido suficiente atención «por separado», en un proceso expansivo de acumular factores de idéntica naturaleza, en un intento de ofrecer una explicación del fenómeno del rendimiento. En el momento de analizar los resultados en cuanto al peso predictivo de cada uno de los sumandos considerados, parece darse un claro consenso en admitir la siguiente gradación:

#### Aptitudes > Ambiente > Instrucción

si bien, como se ha señalado con anterioridad, con frecuencia aparecen inadecuadas interpretaciones del poder explicativo de cada uno de los factores y sumandos al no tener en cuenta los presupuestos estadísticos en los que descansa el análisis de regresión. El lector podrá encontrar en el libro de Marjoribanks (1979) amplia información acerca de la utilización de este modelo –en sus múltiples variantes combinatorias de factores ambientales y aptitudinales (consecuente a su postura teórica deja fuera los factores de instrucción)–.

2.– *Modelos mediacionales*. Parten del supuesto de que determinados factores influyen en el rendimiento académico no de una forma netamente directa, sino a través de la mediación ejercida por otros factores *más próximos* en la explicación teórica asumida. El concepto de *distancia causal* es relevante a los mismos, entendida ésta como la mayor o menor proximidad causal de una variable en el logro de un determinado efecto (rendimiento). Desde mi punto de vista los modelos mediacionales pueden ser considerados como modelos interactivos de causalidad unidireccional en los que la técnica del «path analysis», al descomponer los efectos totales de una variable en efectos directos e indirectos nos ofrece información suficiente para determinar las «mediaciones» ejercidas por determinados factores.

Desde la posición teórica del *symbolic interactionism* se han preconizado la utilización de estos modelos. Los estudios sobre la influencia del *clima* (social y

escolar) en los logros del alumno han venido utilizando con asiduidad este tipo de modelos. Así, Dycr (1972) opina que el ambiente (familiar, escolar y comunitario) afecta las actitudes de padres, profesores y compañeros; éstas influyen en la autoperccepción del alumno que es la que definitivamente determina el rendimiento. Por su parte Brookover *et. al.* (1979) consideran que los diferentes *inputs* sociales de la escuela (características sociales del alumnado) influyen en los logros escolares tanto directamente como a través de la influencia mediacional de las características de la estructura social de la propia escuela y del clima social creado.

En un intento de ampliar el panorama de factores influyentes en los logros escolares, y como resultado del análisis de investigaciones parciales previas, exponemos los modelos propuestos por Centra y Potter (1980, fig. 1) y Glasman y Biniaminov (1981, fig. 2). Las siguientes consideraciones ayudarán a su comprensión:

Fig. 1.—Modelo de Centra y Potter (1980, p. 274)

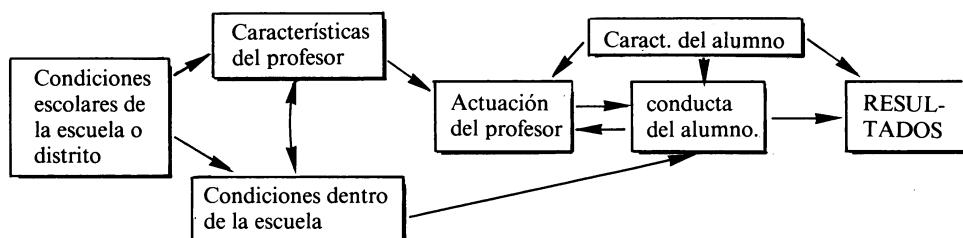
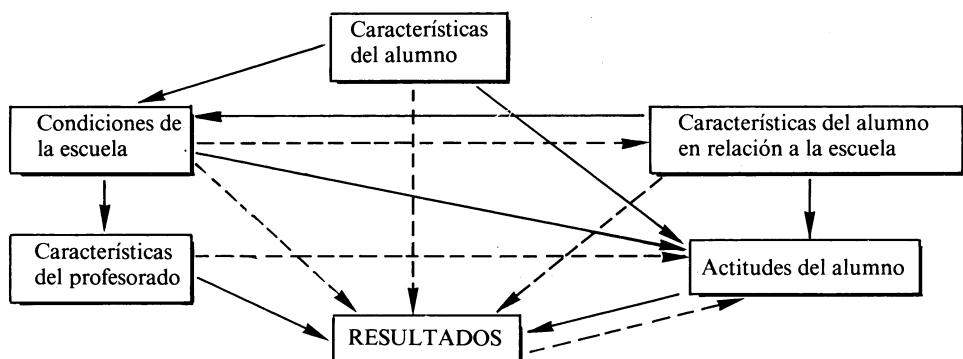


Fig. 2.—Modelo de Glasman y Biniaminov (1981, p. 536)



- En el modelo de Centra y Potter aparece una relación doble de causalidad (verdadera interacción) entre la actuación del profesor y la conducta del

alumno. Esta no recursividad del modelo plantea serios problemas –a mi modo de entender no resueltos– de comprobación estadística que han sido puestos de manifiesto por todos los autores partidarios de los modelos causales. De otra parte, la relación existente entre características del profesor y condiciones dentro de la escuela es meramente asociativa, no presupone ninguna influencia causal.

b) En el modelo de Glasman y Biniamivov las flechas continuas significan una constancia de los hallazgos en investigaciones previas, mientras que las flechas discontinuas ponen de manifiesto inconsistencias en los hallazgos de algunos estudios examinados. El modelo estructural propuesto por dichos autores representa un intento de síntesis globalizadora de los resultados obtenidos en la investigación al utilizar el modelo de *input-output*.

c) Finalmente, hemos de llamar la atención acerca de la dificultad de comprobación que implican modelos globales como los presentados aquí, ya que cada uno de los bloques de factores pueden presentar multitud de indicadores (variables observables). Una vez más se presenta el dilema de elegir para su comprobación aspectos periciales del modelo o reducir el número de indicadores para cada uno de los bloques con el consiguiente agravamiento del problema de los errores de medida. Uno de los criterios sugeridos por numerosos autores es de poner la atención en aquellos factores de posible «manipulación» a través del tratamiento educativo, única vía de alteración de los resultados educativos.

3.- *Modelos interactivos*. Desde el punto de vista teórico fácilmente se asume la «simultaneidad de efectos» en el fenómeno del rendimiento académico. Es decir, los resultados de la enseñanza a la vez que influenciados por una serie de factores (personales, situacionales o didácticos), generan también interacciones recíprocas que afectan a esos mismos factores considerados como determinantes. Un caso de especial atención es el referido a la influencia y efectos circulares de la autoimagen académica y el éxito o fracaso escolar.

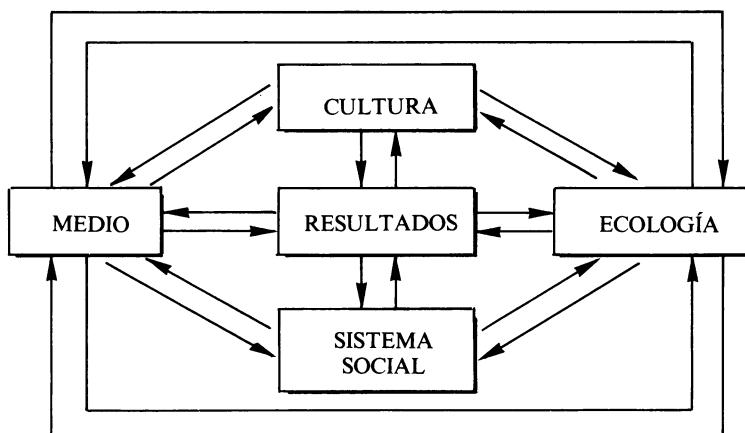


Fig. 3 (Modelo de Anderson)

Sin duda alguna han sido los planteamientos sociológicos y ecológicos sobre la influencia del clima escolar en los resultados de aprendizaje los que han criticado con mayor énfasis el carácter unidireccional de las mediaciones o causalidades atribuidas a una serie de factores ambientales. Anderson (1982, p. 385) recoge una serie de testimonios en esta dirección y, apoyándose en la taxonomía de Tagiuri propone el siguiente modelo teórico (fig. 3) de conceptualización de todas las posibles interacciones entre el clima y los resultados de la enseñanza.

Es obvio afirmar que cualquier modelo interaccionista bidireccional refleja mejor la realidad que los mecanicistas y estáticos análisis unidireccionales de los modelos estructurales, pero la enorme complejidad de los primeros, como señala Saris y Stronkhorst (1984), hace imposible su explicación verbal ni gráfica en un modelo estructural de análisis.

En definitiva, y como afirma Rogosa (1979, p. 298) no existe en el momento presente ningún procedimiento o medida definitiva para la comprobación de los efectos recíprocos, si bien pueden utilizarse algunos indicadores tanto por el método de «cross-lagged correlation» como por el método de ecuaciones estructurales (punto que abordaremos en la parte final de este trabajo).

En resumen, desde el punto de vista operativo y dado el estado actual de las investigaciones sobre el tema del rendimiento, no es pragmática la distinción entre modelos mediacionales e interactivos ya que ambos han de asumir el mismo tipo de explicitación de relaciones entre variables si quieren proceder a una constatación empírica de las mismas. De aquí, que la alternativa de mayor validez y aportación significativa sea la de afrontar investigaciones que se apoyen en un diseño longitudinal-causal, dadas las posibilidades técnicas de análisis que nos ofrecen los programas de la serie LISREL.

## ALGUNAS CONSIDERACIONES SOBRE LOS MODELOS LONGITUDINALES-CAUSALES EN LA INVESTIGACIÓN SOBRE EL RENDIMIENTO ACADÉMICO.

No es el momento de hacer una detallada exposición introductoria sobre el planteamiento, fundamentación y descripción de los estudios longitudinales. La consulta de las obras de Nesselroade y Baltes (eds. 1979) y de Mednick *et. al.* (eds. 1984) son de gran interés tanto desde el punto de vista metodológico como del análisis del contenido y resultados de algunos estudios sobre el tema del rendimiento. Igualmente podemos decirlo de los planteamientos de los modelos causales. En un anterior trabajo (Mateo y Rodríguez, 1984) presentamos una breve introducción al tema así como las referencias básicas para ampliarlo.

La obra de Saris y Stronkhorst (1984) es una valiosa guía para la formulación e interpretación de modelos causales desde la óptica del programa de análisis LISREL.

Nuestra atención se centrará exclusivamente en tres aspectos. En primer lugar en la diferencia de estimación de la relación causal en los diseños longitudi-

nales 2W2V (dos momentos evolutivos –«olas»– dos variables) por el método de *cross-lagged correlation* y el método de ecuaciones estructurales con variables observables. En segundo lugar, el problema de los errores de medida en la utilización de modelos con variables latentes y, finalmente, el problema de la determinación de los efectos recíprocos.

1.- **El diseño de panel longitudinal 2W2V.** Este es el más simple de los diseños longitudinales. En él, un mismo número de casos es observado en dos momentos evolutivos diferentes (no tomamos en consideración la problemática derivada de los efectos de la regresión sobre la media que afecta a toda situación pre-test-post-test). La representación gráfica de este diseño en las dos planteamientos metodológicos que tomamos en consideración sería la siguiente (fig. 4 y 5):

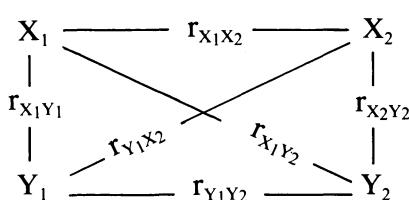


fig. 4. Diseño 2W2V relevante  
al método de *cross-lagged  
correlación*

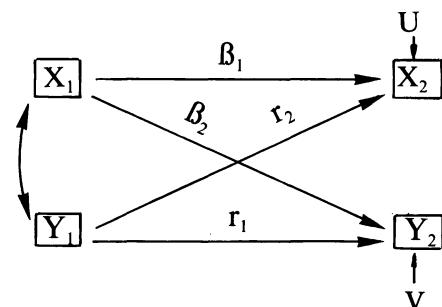


fig.5. Diseño 2W2V relevante  
al modelo causal

Como es conocido, la determinación de la dirección de la causalidad por método de *cross-lagged correlation* viene dada por las diferencias entre las correlaciones de  $X_1$   $Y_2$  e  $Y_1$   $X_2$ . Si  $r_{x_1y_2} > r_{y_1x_2}$ , entonces  $X$  es causa de  $Y$ . Por el contrario, si  $r_{y_1x_2} > r_{x_1y_2}$ , entonces  $Y$  es causa de  $X$ . Siempre en el supuesto de una diferencia estadísticamente significativa. Este método, con sus derivaciones de mayor complejidad es apoyado por autores como Kenny, Calsyn, Crano, Stubbs y otros, y criticado por Duncan, Goldberg, Heise, etc. (vid. Rogosa, 1979). La crítica se centra fundamentalmente en dos puntos. Uno, las diferencias de estabilidad de  $X$  e  $Y$  (indicadas por posible diferencia entre  $r_{x_1x_2}$  y  $r_{y_1y_2}$ ) son consideradas de modo inadecuado por el método al favorecer la dirección causal de la menos estable. En segundo lugar, tampoco se toman en consideración las varianzas de  $X$  e  $Y$  sobre el tiempo (10). En consecuencia, podemos encontrarnos que la diferencia entre los parámetros estructurales *beta* y *gamma* ( $\beta_2 - \gamma_2$ ) sea diferente tanto en signo como en magnitud a la obtenida por el método de *cross-lagged correlation*. Así sucedió al proceder al reanálisis de algunas investigaciones previas.

En consecuencia, y con la precisión de que el modelo causal de la fig. 4 presupone no simultánea causación, medida de las variables sin error y residuos no

correlacionados, creo que hay suficientes argumentos para inclinarse por la utilización del modelo causal en el más simple de los diseños longitudinales.

## 2.- Los modelos causales-longitudinales con variables medidas con error.

Siguiendo en el contexto de los diseños longitudinales 2W2V, uno de los problemas con el que nos encontramos es el del error de medida de las variables observadas. En el modelo presentado en la fig. 6 se presupone que la influencia causal se transmite a través de las variables latentes  $X$  e  $Y$  cuyos indicadores (variables observables y medidas) son  $X_1$  y  $Y_1$  en el primer momento y  $X_2$  e  $Y_2$  en el segundo momento. El problema de este modelo es el de no estar adecuadamente identificado (Saris et. al. 1984) ya que no se da suficiente información para determinar todos los parámetros estructurales.

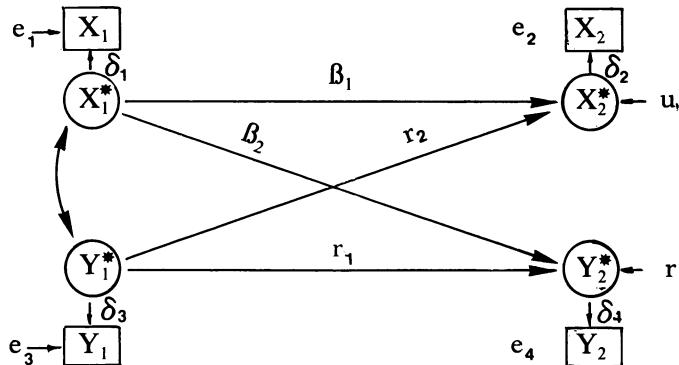


Fig. 6. Modelo causal del diseño longitudinal 2W2V11 con error de medida.

Dos alternativas se sugieren para obviar el problema de la falta de identificación del modelo. Una, la de conocer la fiabilidad (estabilidad) de las medidas de  $X$  e  $Y$  y que, introducidas en el modejo, servirán para obtener una consistente estimación de los parámetros necesarios. La segunda, incrementar el número de medidas de cada variable latente en cada momento evolutivo. Es decir, incrementar el número de variables observadas y medidas para cada constructo hipotetizado. En el modelo de la fig. 7 (2W2V2I) se ejemplifica la duplicación de indicadores para el modelo anteriormente señalado (fig. 6).

Ahora bien, los defensores del método de *cross-lagged correlation* aducen el problema del error de medida como el más grave que afecta al método de ecuaciones estructurales ya que los coeficientes de regresión parciales están más afectados por dicho error que los coeficientes de correlación de orden cero. Por ello, autores como Kenmy (1979) han desarrollado procedimientos para corregir los coeficientes de correlación y controlar el cambio de fiabilidad sobre el tiempo evolutivo considerado.

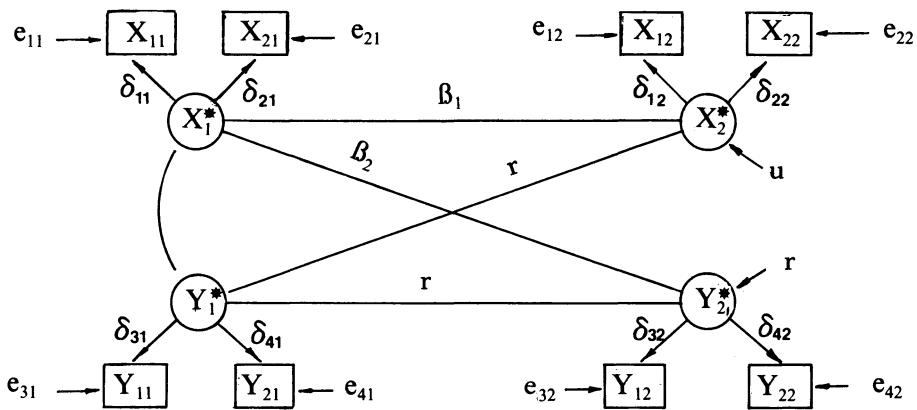


Fig. 7. Modelo causal del diseño longitudinal 2W2V2I que permite la estimación de los parámetros  $\beta_1$ ,  $\beta_2$ ,  $\gamma_1$ ,  $\gamma_2$

La respuesta de los defensores del método de ecuaciones estructurales es la de admitir que tal control puede ejercerse estadísticamente cuando se tiene un sólo indicador por cada variable; sin embargo, afirman que no es posible realizarlo si se tiene información de varios indicadores simultáneos ya que cada «fiable» indicador es tratado separadamente al hacer las comparaciones derivadas del método *cross-lagged correlation* y no de modo conjunto como se realiza en el método de ecuaciones estructurales. Pensemos que el número de comparaciones a realizar viene determinado por  $pq$  ( $\xi$ ), donde  $p$  representa el número de indicadores de una variable,  $q$  los indicadores de la otra variable y  $k$  el número de momentos evolutivos considerados.

En resumen, el problema de los errores de medida no tiene adecuada solución y su magnitud afectará a cualquier modelo, pero creo que puede afirmarse que el método de *cross-lagged correlation* sólo presenta ventajas en el diseño 2W2V1I con las consiguientes limitaciones que representan el tener sólo un indicador de cada uno de los constructos considerados.

**3.- Indicadores de la potencia de los efectos recíprocos.** Tal y como hemos señalado con anterioridad, y pese a la amplia discusión sobre el tema de la estimación de los efectos reciprocos en los estudios longitudinales, la opinión más generalizada es la de que sólo se puede aproximar tal estimación. Por parte del método de *cross-lagged correlation* las medidas descriptivas de la magnitud de las propias *cross-lagged correlations* pueden ser indicadores de tales efectos reciprocos, así como los promedios ponderados de la correlación dentro del momento (*within-time*) y entre variables (*between-variables*) en los diseños comentados de 2W2V.

Por su parte, los modelos causales proponen como indicador de la medida relativa de la potencia de los efectos reciprocos la pérdida de ajuste (*fit*) del modelo restringido con respecto al modelo global. Así, si en el modelo presentación

es la fig. 5 hacemos  $\beta_2 = 0$  y  $\gamma = 0$ ; es decir, no tomamos en consideración dichas influencias causales (*paths*), y comprobamos que los valores de  $\gamma^2$  (Jöreskog & Sorbom, 1977) no sufren significativa alteración (siempre que en sí mismos sean significativos de adecuada ajuste) con respecto al modelo global ( $\beta_2$  e  $\gamma_2 \neq 0$ ), podremos concluir la inexistencia de efectos recíprocos.

Este procedimiento parece más adecuado que el de hipótesis nula de espuriosidad propuesta por Kenny (Rogosa, 1979, p. 299) ya que aún en el supuesto de que se diera que  $\beta_2 = \gamma_2 = 0$ , las diferencias de las *cross-lagged correlations* pueden ser grandes si las estabilidades de X e Y difieren apreciablemente. Sin embargo, la generalización de este procedimiento a diseños de más de dos momentos (*multiwave*) necesita mayor profundización metodológica, aconsejándose «ir partiendo» la secuencia temporal ya que puede ocurrir que la interacción recíproca no sea constante a lo largo de todo el periodo considerado.

## CONCLUSIÓN FINAL

A lo largo de estas páginas muchos interrogantes podrán plantearse, innumerables lagunas quedan, pero después de quince años de preocupación por el tema (en 1970 iniciaba mi tesis de licenciatura sobre «El método predictivo en orientación escolar») ha llegado a la conclusión –siempre provisional– de que el problema central en la investigación sobre el rendimiento académico (dejando aparte el criterio de determinación del mismo) es de carácter metodológico. Tanto por lo que respecta a su evaluación y medida como a la adecuación de los distintos modelos asumidos y, en consecuencia, a la correcta utilización e interpretación de las técnicas estadísticas de análisis. La etapa de desarrollo y aplicación de los modelos causales estructurales abre un camino de grandes perspectivas, a la vez que nos exige un esfuerzo continuado para ofrecer la evidencia suficiente que demuestre que la utilización de los mismos en la investigación educativa permite una mejor comprensión de los factores determinantes del rendimiento académico. Dicha comprensión debe constituir la base para la puesta en marcha de adecuadas y efectivas intervenciones educativas.

## NOTAS

1. Especial atención merecen los cuatro informes americanos: Adler, M. (1982). *The paideia proposal: An educational manifesto*. New York: MacMillan.
- Boyer, E. (1983). *High school: A report on secondary education in America*. New York: Harper & Row.
- Gardner, D. (1983). *A nation at risk: The imperative for educational reform*. Washington, DC: U.S. Department of Education.
- Goodlad, J. (1983). *A place called school: Prospects for the future*. New York: McGraw-Hill.
- Un análisis de los mismos puede encontrarse en: Aubrey, R.F. «Reform in Schooling: Four Proposals on an Educational Quest». *Journal of Counseling and Development*, 1984, 63, 204-213.

2. Un punto de vista positivo se desprenden de la investigación de Rutter, M., Maughan, B., Mortimore, P., Ouston, J. (1979) *Fifteen Thousand Hours-Secondary Schools and their Effects on Schoolchildren*. London: Open Books.

3. El denominado «Wisconsin Path Model» (Marjoribanks, 1979, p. 105), de gran influencia en la investigación llevada a cabo por los sociólogos americanos, considera tres indicadores del «status social» relacionados según la siguiente cadena causal:

→ Logros educativos → Logros profesionales → Salario, considerando como bloques de factores causales previos: «Background» social → Desarrollo y rendimiento intelectual → Influencias significativas de «otros» → Expectativas del alumno.

4. En estos momentos se asiste en Gran Bretaña a un interesante polémica acerca del papel de control de las autoridades centrales sobre las certificaciones académicas. La reflexión sobre la dialéctica autonomía educativa y curricular vs. control de la calidad del producto educativo final debiera ser objeto de especial consideración en los planteamientos de nuestra reforma educativa.

5. Respecto a este rendimiento mediato, Rutter *et. al.* (1980, pg. 171) afirman que las escuelas son fundamentalmente diseñadas para lograr objetivos educativos y sería bastante inapropiado expresar sus metas en términos de ocupaciones o ingresos alcanzados por sus alumnos. Una vez más aparece la polémica sobre la función de cambio social de la educación.

6. J. Garanto, J. Mateo y S. Rodríguez: «Modelos y técnicas de análisis del rendimiento académico» *Revista de Educación* (En prensa)

7. Pueden examinarse los trabajos de Bangert *et. al.* (1983) y de Arlin (1984) sobre las críticas a las investigaciones de Bloom y colaboradores, y en general a los hallazgos sobre los sistemas de instrucción en la escuela secundaria. Scheirer y Kraut (1979) analizan los efectos de determinados programas sobre el incremento del autoconcepto.

8. Del estudio de E. Schiefelbein y J. Simmons: «Los determinantes del rendimiento escolar. Reseña de la investigación para los países en desarrollo», *Educación hoy*, 1980, 60 (monográfico), hemos extraído un «decálogo» de hallazgos significativamente constantes en estudios llevados a cabo en un periodo de 20 años en países en desarrollo. Las implicaciones de dicho decálogo son obvias en la toma de decisiones sobre política educativa.

1. El tamaño mayor de la clase se asoció con un mejor aprovechamiento o no lo afectó.
2. El mayor gasto por estudiante no está asociado a mayor rendimiento.
3. Los alumnos que tienen deberes fuera del aula rinden más
4. Los maestros sin título de capacitación docente tenían alumnos con resultados tan buenos como aquellos que poseían el título
5. Los años de experiencia del maestro aparecen como un determinante de interés en sólo 7 de los 19 estudios analizados.
6. Un mayor tiempo de capacitación del maestro no produce mejores resultados en los alumnos.
7. Los antecedentes familiares de los alumnos son el determinante individual de mayor importancia en los resultados escolares.
8. A mayor número de repeticiones del alumno, menor rendimiento.
9. La asistencia en preescolar tiene gran influencia en el rendimiento posterior.
10. Desnutrición, peso corporal y salud física fueron predictorés significativos del rendimiento académico.

9. Athappilly, K.K. (1978) «a meta-analysis of the effects of modern mathematics in comparison with traditional mathematics in the American educational system» (Doct. dissert., Western Michigan University). *Dissertation Abstracts International*, 1979, 39, 6002 (University Microfilms n.º 79-08643)

- Baker, S.B., Swisher, J.D., Nadenichek, P.E., Popowicw, C.L. (1984) «Measured effects of primary prevention strategies», *Personnel & Guidance Journal*, 62, 459-464.
- Bangert, R.L., Kulik, J.A., Kulik, Chen-Lin C. (1983) «Individualized systems of instruction in secondary Schools», *Review of Educational Research*, 53, 143-158.
- Bredderman, T. (1978) *Elementary school process curricula - A meta-analysis*. Albany, State University of New York (ERIC n.º ED 170333)
- Cohen, P.A., Kulik, J.A., Kulik, C-L.C. (1982) «Educational outcomes of tutoring: A meta-analysis of findings», *American Educational Research Journal*, 19, 237-248.
- Glass, G.V., Smith, M.L. (1978) *Meta-analysis of the research on the relationship of class-size and achievement*. S. Francisco: Far West Laboratory for Educational Research and Development.
- Hansford, B.C., Hattie, J.A. (1982) «The relationship between self and achievement/performance measures», *Review of Educational Research*, 52, 123-142.
- Hartley, S.S. (1977) «Meta-analysis of the effects of individually paced instruction in mathematics» (Doct. Dissert, University of Colorado) *Dissertation Abstracts International*, 1978, 38 (7-A), 4003. (University Microfilm n.º 77-29926)
- Kulik, J.A., Kulik, C-L.C., Cohen, P.A. (1979) «A meta-analysis of outcome studies of Keller's personalized system of instruction», *American Psychologist*, 38, 307-318
- (1979) «Research on audio-tutorial instruction: A meta-analysis of comparative studies», *Research in Higher Education*, 11, 321-340
- Kulik, C-L.C., Kulik, J.A. (1982) «Effects of ability grouping on secondary of advanced organizers on learning and retention», *American Educational Research Journal*, 17, 211-218
- Smith, M.L., Glass, G.V. (1979) *Relationship of class-size to classroom processes, teacher satisfaction and pupil affect: A meta-analysis*. S. Francisco: Far West Laboratory for Educational Research and Development.
- Steinkamp, M.V., Maehr, M.L. (1983) «Affect, ability, and science achievement: A quantitative synthesis of correlational research», *Review of Educational Research*, 53, 369, 396.

10. La razón técnica es expresada por Ragosa (1979, pg. 278) en virtud de la relación existente entre las *cross-lagged correlations* y los parámetros estructurales del modelo causal.

$$[\ell_{X_1Y_2} - \ell_{Y_1X_2} = (I - \ell^2 X_1 Y_1) [\frac{\beta_2(\sigma_{X_1})}{\sigma_{Y_2}} - r_2(\sigma_{Y_1})] + \ell_{X_1Y_1} (\ell_{Y_1Y_2} - \ell_{X_1X_2})]$$

#### REFERENCIAS BIBLIOGRÁFICAS

- ACTON, T.A. (1980). «Educational criteria of success: some problems in the work of Rutter, Mauhan, Mortimore and Ouston». *Educational Research*, 22, 163-169.
- ANDERSON, C.S. (1982). «The search for school climates: A review of the research». *Review of Educational Research*, 52, 368-420.
- ARLIN, M. (1984). «Time, equality, and mastery learning». *Review of Educational Research*, 54, 65-86
- ASTIN, A.W. (1984). «Using longitudinal data to study College impact». En S.A. Mednick, M. Harway and K.M. Finello (Eds). *Handbook of Longitudinal Research*. V. II (pp. 62-76). Praeger Pub. (2 v.). New York.
- BANGERT, R.L., Kulik, J.A., y Kulik, Ch. C-L., (1983) «Individualized systems of instruction in secondary Schools». *Review of Educational Research*, 53, 143-158.
- BERNSTEIN, B. (1970). «Education cannot compensate for society». *New Society*, 387.

- BLOOM, B.S. (1964). *Stability and change in human characteristics*. Wiley. New York.
- BRIDGE, R.G., JUDD, C.M., y MOOCK, P.R. (1979). *The determinants of educational outcomes*. Ballinger. Cambridge, Mass.
- BROOKOVER, W.B., BEADY, C., FLOOD, P., SCHWEITZER, J., y WISENBAKER, J., (1979). *School social systems and student achievement*. New York: Praeger Pub.
- CATTEL, R.B., y BUTCHER, H.J. (1968). *The prediction of achievement and creativity*. Bobbs-Merrill. New York.
- CENTRA, J.A. y POTTER, D.A. (1980). «School and teachers effects: An interrelational model». *Review of Educational Research*. 50, 273-291.
- COLEMAN, J.S., CAMPBELL, E.Q., HOBSON, C.J., McPARLAND, J, MOOD, A., WEINFELD, F.D., y YORK, R.L., (Eds) (1966). *Equality of educational opportunity* (2 v) U.S. Government Printing Office. Washington D.C.
- DOYLE, W. (1983). «Academic work». *Review of Educational Research*. 53 159-199.
- DYER, H.S. (1972) «Some thoughts about future studies». En F. Mosteller, y D.P. Moynihan (eds). *On equality of educational opportunity*. Randon House. New York.
- GLASMAN, N.S. y BINIAMINOV, I. (1981) «Input-Output analyses of schools.» *Review of Edcuational Research*. 51, 509-539
- GRAY, J. (1981). «School effectiveness research Key issues». *Educational Research*, 24, 49-54.
- HOGE, R.D. y LUCE, S. (1979). «Predicting academic achievement from classroom behaviour». *Review of Educational Research*, 49, 479-499.
- JENKS, C.S., SMITH, M., ACLAUD, H., BANE, M.J., COHEN, D., GINTIS, H., HEYNS, B., y MICHELSON, S. (Eds) (1972). *Inequality: A reassessment of the effect of family and schooling in America*. Basic Books. New York.
- JÖRESKOG, K.G. y SÖRBOM, D. (1977) «Statistical Models and methods for the analysis of longitudinal data» En: D.J. Aigner y A.S. Goldberger (Eds): *Latent Variables in Socio-economic models*. (pp. 285-326); North Holland Pub. Co. Amsterdam.
- KEEVES, J.P. (1972). *Educational environment and student achievement*. Almquist & Witsell. Stoc-kholm.
- KENNY, D.A. (1979). *Correlation and causality*. Wiley. New York.
- KURPIUS, D.J. (Ed) (1978). *Learning: Making learning environments more effective*. Accelerated Development. Muncie.
- LAVIN, D.E. (1965). *The prediction of academic performance*. Rusell Sage Foundation. New York.
- LUECKE, D.E. y McGINN, N. (1975). «Regression analysis and education production functions: Can they be trusted? *Harvard Educational Review*, 45, 325-350.
- MADAUS, G.F., KELLAGHAN, TH., RAKOW, E.A. y KING, D.J. (1979). «The sensivity of me-sures of school effectiveness. *Harvard Educational Review*, 49, 207-230.

- MARJORIBANKS, K. (1974) (Ed) *Environments for learning*. NFER. Slough.
- MARJORIBANKS, K. (1979). *Families and their learning environments: An empirical analysis*. Routledge & Kegan Paul. London.
- MATEO, J. (1980). *Factores de personalidad en el rendimiento escolar*. Tesis doctoral inédita. Universidad de Barcelona.
- MATEO, J y RODRÍGUEZ, S. (1984) «Precisiones y limitaciones explicativas en los métodos correlacionales. Alternativas metodológicas». *Revista de Investigación Educativa*, v. 2. 4, 103-132.
- MAYESKE, G.W. et al. (1972). *A study of our nation's schools*. U.S. Government Printing Office. Washington D.C.
- MEDNICK, S.A., HARWAY, M. y FINELLO, K.M. (Eds) (1984) *Handbook of longitudinal research* (2 v) Praeger Pub. New York.
- MOOS, R.H. (1979). *Evaluating educational environments*. Jossey-Bass. San Francisco.
- PELECHANO, V. (1977). *Personalidad, inteligencia, motivación y rendimiento en el B.U.P.* ICE Universidad de La Laguna. La Laguna:
- RODRÍGUEZ, S. (1982). *Factores de rendimiento escolar*. Oikos-Tau. Barcelona.
- ROGOSA, D. (1979). «Causal models in longitudinal research: rationale formulation and interpretation» En: J.R. NESSELROADE y P.B. BALTES (Eds) *Longitudinal research in the study of behaviour and development*. 263-302. Academic Press. New York.
- RUTTER, M., MAUGHAN, B., MORTIMORC P. & OUSTON, J., (1979). *Fifteen thousand hours-secondary schools and their effects on schoolchildren*. Open Books. London.
- RUTTER, M., MAUGHAN, B., MORTIMORE, P., y OUTON, J., (1980). «Educational criteria of success: A reply to Acton». *Educational Research*, 22, 170-174.
- SARIS, W y STRNKHORST, H. (1984) *Causal modelling in nonexperimental research*. Sociometric Research Foundation. Amsterdam.
- SCHEIRER, M.A. y KRAUT, R.E. (1979) «Increasing educational achievement via self-concept changes». *Review of Educational Research*, 49, 131-150.
- SCHIEFELBEIN, E y FARRELL, J. P. (1982) *Eight years of their lives*. International Development Research Centre. Ottawa.
- SHULMAN, L.S. (Ed) (1977) *Review of Research in Education* (v. 5). F.E. Peacock. Itasca IL.
- SNOW, R.E. (1973) «Theory construction for research on teaching» En R.M.W. Travers (Ed). *Second handbook of research on teaching*. Rand McNally. Chicago.
- WALBERG, H.J. (1974). «Models for optimizing and individualizing school learning». En K. Marjoribanks (Ed.) *Environment for learning* (pp. 19-38). NFER. Slough: (Publicado originalmente en *Interchange*, 1971 3, 15-17.
- WALBERG, H.J. (Ed.) (1979). *Educational environments and effects*. McCutchan. Berkeley.
- WASNA, M. (1974). *La motivación, la inteligencia y el éxito en el aprendizaje*. Kapelusz. B. Aires.

## **Evaluación del rendimiento criterial versus normativa. Modelo de evaluación: F.C.O.**

*por  
Mercedes Rodríguez Lajo*

### **INTRODUCCIÓN**

La presente ponencia está estructurada en dos partes.

En la primera de ellas se esboza un análisis entre comparativo los tests referidos al criterio y los tests referidos a la norma, haciendo más énfasis en los primeros. Se inicia con un breve resumen histórico de cómo surgen éstos, ante una necesidad de más riqueza en las informaciones que las proporcionadas por los instrumentos normativos. A continuación se analizan las diferencias entre ambos tipos de instrumentos para finalizar con unas anotaciones sobre la problemática que plantean los tests referidos al criterio en cuanto a la validez y fiabilidad se refiere.

Finalizada esta parte, que equivale a una síntesis teórica, un resumen esquemático del proceso de evaluación introduce a la segunda, que es la descripción de un modelo de evaluación criterial desarrollado por la ponente.

El modelo, denominado F.C.O., está construido siguiendo los pasos de un modelo trifásico de evaluación del rendimiento (preparación-implementación-interpretación) y su interés estriba en que contempla integralmente todos los aspectos del mismo, pudiendo, en consecuencia, ser una herramienta muy eficaz para el control total de un curso académico. A lo anterior coayuda el hecho de que el modelo está totalmente informatizado en un ordenador personal.

En el anexo se facilitan unos gráficos de control proporcionados por el modelo, de sumo interés en el mismo.

#### **1. Surgimiento de la evaluación criterial frente a la normativa.**

La evaluación criterial nace como consecuencia de una convergencia de hechos acaecidos en el cambio de década 1950-60: difusión de la enseñanza «pro-

gramada», el desarrollo de las teorías del aprendizaje y el auge de los objetivos conductuales y el enfoque de la enseñanza basada en objetivos. Se crea así un clima social que pone de manifiesto la necesidad de nuevos instrumentos que aporten una mayor información sobre lo que «un sujeto es capaz de hacer». Estos son denominados «tests referidos al criterio» (TRC), término propuesto por vez primera por Glaser, R. en 1963, y que define cómo: «Los que dependen del estatus absoluto de cualidades del estudiante, mientras que un test orientado a las normas dependen del estatus relativo.»

La asunción de la normalidad hacia que los instrumentos existentes se elaboraran y diseñaran de forma tal que permitieran establecer diferencias entre los sujetos comparándolos entre sí. En los tests así construidos, (tests referidos a la norma -TRN-) la puntuación directa da una información global de ejecución del sujeto y las puntuaciones se interpretan en función de la ejecución del grupo.

Ya Thorndike en 1913 señalaba cómo un sistema de calificaciones comparativas no proporcionaba una imagen clara de las habilidades del sujeto. Popham, W.J. (1978), en la misma línea, añade que «su excesiva generalidad lleva a desfases entre lo que se evalúa y lo que se enseña; el no proporcionar pistas suficientes para poner remedio a programas educativos y el emplear procedimientos técnicos en la construcción de instrumentos que tienden a eliminar los ítems más relevantes».

Desde su nacimiento, los TRC despiertan un gran interés, proliferando rápidamente los escritos sobre ellos, señalándose ya en 1978 por Hambleton, R.K. y colaboradores más de 600 referencias bibliográficas. Ello produce un lógico confusionismo terminológico. Así Gray (1978) señala 57 definiciones distintas en 40 autores, demostrando que no sólo autores distintos utilizan distintos términos sino que incluso algunos autores son inconsistentes en un mismo artículo.

Así (Glass, G. 1978), se identifican «criterio» y «estandar» o luego «criterio» y «objetivo conductual», conceptos que posteriormente aparecen claramente diferenciados. Nacen expresiones no exactamente equiparables: «tests referidos a objetivos», «de competencia mínimas», «de destrezas básicas», «de dominio», etc.

Nitko A.J. (1980) tras analizar todos los instrumentos elaborados hasta 1978 llega a la conclusión de que no hay un «prototipo» de TRC, sino que es un concepto que ha emergido con múltiples facetas pero con un propósito común: *«el emplear como referente para la interpretación de puntuaciones un dominio que debe estar previamente bien definido».*

Destacamos, finalmente, las aportaciones que hace Hambleton, R.K. (1980) en la línea de la definición anterior.

- Términos como «objetivos», «competencias» y «destrezas», pueden usarse indistintamente.
- La definición del dominio no tiene porqué incluir una referencia respecto a una puntuación de corte (cut-off) o estándar.
- Cuando se mide más de un objetivo en un test, los ítems del mismo están organizados en subtests, que se corresponden con cada objetivo.

A la síntesis señalada de las conceptualizaciones de Nitko y Hambleton, con las que estamos de acuerdo, añadiremos que: el concepto de «criterio» en los TRC se refiere a un dominio de contenido y de comportamiento hacia el cual se refiere la puntuación del test.

## 2. Diferencias entre los T.R.C. y los T.R.N.

La mayoría de autores que describen las diferencias entre estos dos tipos de instrumentos (Block, J.H. 1971; Glaser, R. y Nitko, A.J. 1971; Popham, W.J. y Husek, T.R. 1969) coinciden en afirmar que no se distinguen a simple vista.

Las diferencias se sitúan en torno a estas tres áreas:

### *A. Especificación del contenido de dominio medido por el test.*

En los T.R.C. el dominio tiene que ser descrito de forma mucho más exhaustiva. Se pretende llegar a una descripción del mismo que sea lo suficientemente breve para resultar útil y por otra, que circunscriba y delimite suficientemente la clase de conductas a estudiar, para reducir ambigüedades. Para ello se elaboran una serie de estrategias, que a su vez facilitarán, posteriormente, el muestreo y la selección de ítems, como: «los objetivos de comportamiento, «las formas de ítems» (Hively, W.B.; Patterson H.L. y Page, S.A. 1968), «los objetivos ampliados» (Millman, J. 1974), «las especificaciones del test IOX (Popham, W.J. 1983) etc.

### *B. Desarrollo, análisis y selección de ítems.*

En los T.R.N. los ítems:

- \* Se construyen en función de objetivos.
- \* Se analizan y seleccionan no sólo en función de su adecuación al objetivo sino también de sus características estadísticas, eliminando los no discriminativos. Predominan procedimientos empíricos «a posteriori».

En los T.R.C.,

- \* Se construyen a partir del universo de dominio «bien definido».
- \* Se analizan y seleccionan las agrupaciones de ítem/objetivo («clusters») buscando que los mismos sean representativos. Predominan los procedimientos «a priori» de análisis y se utilizan métodos de muestreo en la selección, según las características del dominio.

### *C. Interpretación de las puntuaciones.*

Si, como afirma Tenbrink, T.D. (1981), un referente es «aquellos a lo que uno se refiere cuando formula un juicio» en los T.R.N. el referente es relati-

vo (la ejecución del grupo) y en los T.R.C. el referente es absoluto («el dominio»).

Otra diferencia, señalada por Mehrens, W. (1982) es que en los T.R.N. el número de objetivos muestreados suele ser más amplio que en los T.R.C. Estas pruebas cubren menor ámbito de generalidad.

Las diferencias señaladas parecen indicar que los T.R.N. sirven mejor a propósitos sumativos mientras que los T.R.C. sirven mejor a propósitos formativos. Paralelismo que no tiene porque darse. Si bien es cierto que para propósitos formativos es necesaria la referencia criterial, la información normativa también será útil y, viceversa, la referencia criterial será de sumo interés en los propósitos sumativos. Toda prueba se construye de acuerdo con los objetivos y la riqueza informativa que de ella se obtenga dependerá de la rigurosidad con que se ha construido, en cuanto a su representatividad. Sin representatividad no podrá realizarse la interpretación criterial, pero sí la normativa.

### **3. La validez y fiabilidad en las pruebas de referencia criterial: problemas que plantean.**

Dadas las características de estos instrumentos se presenta una serie de problemas que afectan directamente a su valoración.

*La Fiabilidad* no es una propiedad intrínseca de los instrumentos, sino de éstos cuando se aplican, se refiere a los resultados obtenidos. Será una condición necesaria para la validez, pero no suficiente.

En torno a la fiabilidad se presentan dos problemas:

#### *1) La reducción de la variabilidad de las puntuaciones.*

En la Psicometría clásica los índices para determinarla se basan, en general, en el cálculo del coeficiente de correlación de Pearson. Como éste depende de la variabilidad de las puntuaciones, si ésta es baja, también lo será el coeficiente de fiabilidad. Este hecho, constatado por muchos autores, ha llevado a decidir que el método correlacional puede ser engañoso e inadecuado. Habrá que interpretarlos con cautela y si la varianza es reducida, descartarlos.

Prefieren por eso hablar de «índices de acuerdo», más que de «coeficientes de fiabilidad» pues dichos enfoques no están definidos ni interpretados en componentes de varianza.

#### *2) El contexto de la decisión.*

La naturaleza de las alternativas de decisión influye sobre el tipo de procedimientos de fiabilidad que se empleen.

Surgen distintas conceptualizaciones de fiabilidad y procedimientos alternativos para su estimación. Hambleton R.K. y colaboradores (1978) distinguen tres conceptualizaciones de fiabilidad en torno a las cuales se pueden clasificar los métodos.

*a) Fiabilidad como medida de acuerdo en las decisiones de clasificación.*

Se enfatiza en la consistencia de las decisiones cuando se repite un instrumento en un mismo grupo a través de formas paralelas o aleatoriamente paralelas.

Estarían, entre otros, los procedimientos de Hambleton, R.K. y Novick, M.R. (1973); Huynh, H. (1976) o Subkoviack, M.J. (1976). Se denominarán «índices de pérdida de umbral». Implicarán puntuaciones, «cut-off», en función de la cual se tomarán las decisiones y las pérdidas asociadas a los errores de decisión, se consideran de igual importancia.

*b) Fiabilidad de las puntuaciones cut-off.*

Se enfatiza en la consistencia de las desviaciones de las puntuaciones respecto a las puntuaciones cut-off. Estarían, entre otros, los procedimientos derivados de la teoría de la generalización, índices de Brennan, R.L. y Kane, M.T. (1977); Kane, M.T. y Brennan, R.L. (1980) y Livingston, S.A. (1972). Se denominan índices de «pérdida de varianza». Implican una puntuación de corte y las pérdidas asociadas a los errores de decisión no se consideran de igual importancia.

*c) Fiabilidad del dominio del individuo.*

Se enfatiza en la consistencia de las puntuaciones respecto al dominio estimado de los sujetos, sin hacer referencia a las puntuaciones de corte. Entre otros procedimientos estarían los de Lord, F.M. y Novick, M.R. (1968) y Brennan, R.L. (1980).

*La Validez*

Características más importante de los instrumentos que lleva a tener que determinar sobre el mismo desde su mismo objeto de medida: ¿qué mide?, hasta su finalidad: ¿para qué lo mide?, pasando por el: ¿cómo y de qué forma lo mide?; en donde no sólo deben considerarse aspectos organizativos y estructurales del mismo, sino que también dependerá de la fiabilidad de las puntuaciones.

Hay que señalar, sin embargo, que a pesar de la numerosa contribución literaria, la validez de los T.R.C. *no está resuelta de forma satisfactoria*. En ellas seguirá siendo una cuestión importante pero variará el grado de interés puesto en cada uno de los distintos tipos de validez. Se referirá también «a la interpretación de las respuestas» y a «las decisiones» que de ellos se tomen. Veamos de forma muy sintética los distintos tipos de validez en los T.R.C.

#### *La validez de criterio o funcional.*

Será útil cuando queramos saber el grado en que un instrumento covaría con otros que pretendan medir lo mismo y cuando nos interese saber si el instrumento discrimina entre sujetos clasificados sobre criterios diferentes.

Se presentarán los siguientes problemas para su determinación: la validez de los criterios externos que sirven de apoyo, la falta de variabilidad de las puntuaciones y la influencia que sobre la validez ejerce la fiabilidad.

Muchos autores opinan que en el contexto de los T.R.C. es irrelevante, considerando que el instrumento ya, por sí mismo, es un buen criterio.

Se sugerirán distintas técnicas como: los grupos contrastados (Haynes, 1978) o la estadística discriminativa (Mehrens, W. 1982).

#### *La Validez Descriptiva o de Contenido.*

Es lo más importante, considerándose en general, que por sí sola es suficiente para valorar estos instrumentos. Según Berk, R.A. (1980, implicaría:

- a) *La especificación del dominio.*
- b) *La validez de ítems. Existiendo distintos procedimientos: «Judiciales» y «Empíricos» para establecer la congruencia entre los ítems y las especificaciones de dominio.*
- c) *La calidad técnica de los ítems.*
- d) *La representatividad de los mismos para que permitan generalizar.*

Recogemos, a modo de conclusión, las siguientes palabras de Linn, R.L. (1980): «La validez de contenido proporciona una excelente fundamentación para los T.R.C., pero para soportar la validez de un instrumento se necesita, además, la validez de las inferencias y de los usos del mismo». Es decir, es un tipo de *validez necesaria pero no suficiente*.

#### *La Validez de Constructo.*

Ha sido infravalorada en el marco de la evaluación criterial. Atenderá a la *significación de las puntuaciones* del test en relación con su estructura inicial.

Las especificaciones de los objetivos, que se realiza en estos instrumentos, es similar a lo que Kerlinger denomina «definición operativa del constructo».

Martínez, M.R. (1981) destaca este hecho señalando como los constructos son los que determinan qué conductas han de seleccionarse para su observación y la situación en que será aplicada y como todo sistema de medida está siempre ligada al sistema de constructos dentro de cuál fue construido.

A este tipo de validez no se le ha dado la importancia suficiente.

#### *La Validez de Decisión o de Utilidad.*

Es quizás el logro más importante, pero a su vez más controvertido, de los T.R.C. a la validez.

Un instrumento será válido si es útil a los propósitos para los cuales fue elaborado. Si de ellos deben desprenderse decisiones, será importante determinar su adecuación. Se analizará este tipo de validez en función de los distintos contextos de decisión.

Así Cronbach, L.J. (1970) diferencia entre dos niveles: las que atañen a individuos y las que atañen a grupos.

Shepard, L. (1980) diferencia, dentro de las individuales, entre: para propósitos instructivos y para «certificación».

Siempre que se tomen decisiones concernientes a la clasificación habrá que establecer puntuaciones cut-off.

La «fijación de los standars» será el problema fundamental de la metodología de los T.R.C. Su importancia nos la señala Hambleton, R.K. (1982) cuando afirma: «por muy acertado que sea técnicamente un test o por mucha validez de contenido que tenga, todos los esfuerzos serán vanos si no se presta suficiente atención al proceso de establecer standars».

Habrá distintos tipos de standars (para cada objetivo, para el conjunto de objetivos, o para la ejecución del grupo) y distintos procedimientos para establecerlos (Meskauskas, J.A. 1976; Glass, G. 1978 y Martínez, M.R. 1982).

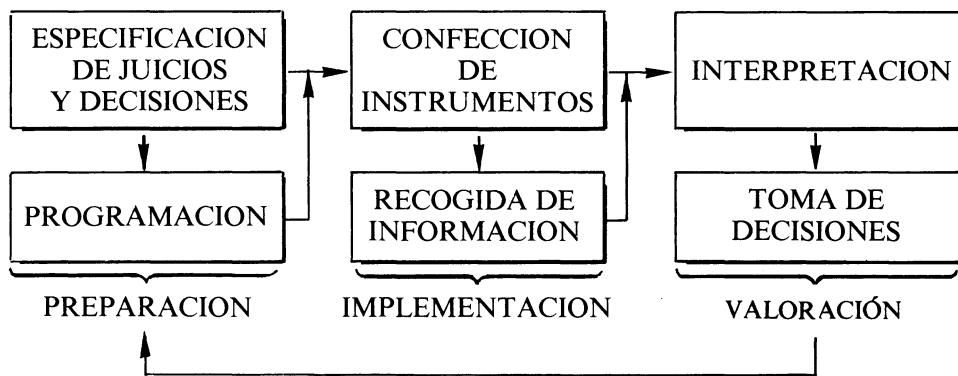
El hecho de que al aplicar distintos métodos a un mismo test que tiene un único propósito, se obtengan standars diferentes ha hecho que se ponga en tela de juicio los métodos. Será un problema que afectará directamente a la validez de la decisión. Lo importante, después, será determinar que las decisiones que se tomen son las correctas.

Los conceptos de validez y fiabilidad se funden en el marco de la Teoría de la Generalización (Cronbach, L.J. y col. 1972).

#### **LA EVALUACIÓN DEL RENDIMIENTO: EL MODELO F.C.O.**

El proceso de evaluación puede ser representado esquemáticamente por el siguiente modelo trifásico, en el cual, como puede verse, todas las fases están concatenadas y relacionadas.

## PROCESO DE EVALUACIÓN. MODELO TRIFASICO



El modelo de evaluación F.C.O. que se presenta en esta ponencia, objeto de mi tesis doctoral, aún en fase elaboración, sigue el modelo trifásico expuesto y siendo eminentemente criterial posee elementos normativos complementarios.

### **El Modelo «F.C.O.»**

A continuación se presenta un modelo de evaluación criterial del rendimiento académico mediante pruebas de elección múltiple.

A dicho modelo le hemos denominado F.C.O., iniciales, respectivamente, de «facilidad» y «cut-off», debido a que la característica más sobresaliente del mismo es, como veremos, que las distintas puntuaciones «cut-off» se establecen automáticamente a partir de las facilidades de los ítems que intervienen en las pruebas.

Una faceta fundamental del modelo que se presenta es que está totalmente informatizado en un ordenador personal. Ello permite al educador la utilización del mismo en su casa, según sus disponibilidades de tiempo.

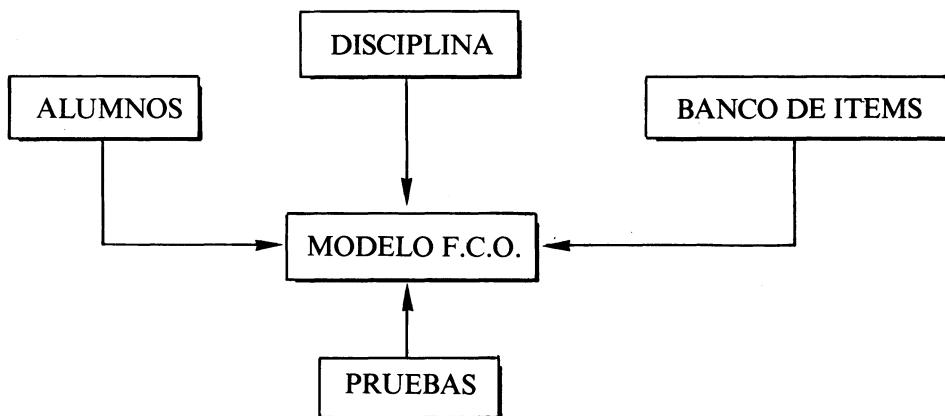
Debido a esta informatización, el modelo, integrado junto con otros procesos, permite el control y seguimiento al día del curso académico del cual es responsable el educador.

### **1. Componentes que intervienen en el modelo**

Dado que el modelo controla el desarrollo de un curso académico veamos a continuación qué componentes son precisos para el control del mismo.

Evidentemente el primer componente es la población de sujetos a los cuales se desea controlar. Otra componente es la disciplina que se enseña, la cual debe

## MODELO QUE CONTROLA UN CURSO ACADEMICO



estar «bien definida» y estructurada coherentemente «a priori» al inicio del curso. Finalmente, son precisas unas pruebas periódicas para efectuar las sucesivas evaluaciones.

Las componentes anteriores son incorporadas en el modelo como archivos, los cuales se van relacionando y actualizando convenientemente. Se incorpora, asimismo, un archivo auxiliar, pero de suma importancia, que es el banco de ítems, el cuál recoge la experiencia y resultados históricos de todas las pruebas realizadas anteriormente.

Las informaciones que recoge cada uno de los archivos anteriores, escuetamente, son las siguientes:

### A. Archivo DISCIPLINA.

Es donde se guarda lo que constituye el «Dominio» de la materia. Esta es considerada como «un todo». El «Dominio» está definido y acotado (cuya adecuación o bondad se debe validar) por los «objetivos específicos»  $\{w_i\}$ . Así la disciplina  $\{D\}$  se equipara a un conjunto de objetivos.

Por otra parte, cada objetivo se articula en «especificaciones operativas» agrupadas en torno a competencias concretas.

Para cada objetivo debe establecerse asimismo un valor escalar denominado «peso» ( $\pi_i$ ) el cual marca la importancia de cada uno de ellos en el dominio de la disciplina  $\{D\}$ . Se obliga en el modelo a que  $\sum \pi_i = 100$ .

Por otra parte, debe determinarse «a priori» si un objetivo es considerado «esencial» o no, ya que los que posean este atributo son evaluados por el modelo más rígidamente, como se verá posteriormente.

El procedimiento seguido para establecer todo lo anterior es a criterio del educador. Sin embargo, nosotros hemos optado por el sistema judicial, existien-

do marginalmente impresos de encuesta a jueces y algoritmos programados para su ejecución.

El archivo resultante se compone, en consecuencia, de las siguientes informaciones:

- Código del Objetivo
- Descripción
- Peso
- ¿Es esencial?
- Número de especificaciones operativas componentes
- Etc.
  
- Número de especificación operativa
- Descripción
- Objetivo a que pertenece
- Facilidad histórica medida de la especificación
- Etc.

## B. Archivo ALUMNOS.

Corresponde a la población de control. Equivale a la ficha individual que se realiza a inicio de curso. Contiene, en consecuencia, de entrada, informaciones tales como:

- Número de alumno
- Nombre
- Dirección
- Sexo
- Experiencia previa en la disciplina
- Etc.

A estas informaciones fijas el modelo actualiza, posteriormente, unas informaciones variables que son las distintas puntuaciones que se van obteniendo a partir de las pruebas efectuadas.

*Tipos de puntuaciones* que se recogen, fruto de la aplicación del modelo:

$q_i$  = puntuación relativa de cada individuo en cada objetivo {0,÷ 10}.

$|q_i|$  = puntuación absoluta de cada individuo en cada objetivo =  $q_i \times \pi_i$ .

$Q$  = Puntuación total de cada individuo en una prueba. Al valor máximo posible que se puede obtener se le denomina  $Q^{\max}$ .

$$Q_{10} = \text{puntuación en base 10 de una prueba} = \frac{Q}{Q^{\max}} \times 10.$$

$Q^D$  = puntuación total del individuo en la disciplina hasta el momento.

Es igual a la suma de las puntuaciones  $Q$  obtenidas anteriormente.

Al valor máximo posible obtenible hasta el momento se le denomina  $Q_{\max}^D$ .

$Q_{10}^D$  = puntuación de base 10 de la disciplina hasta el momento =  $\frac{Q^D}{Q_{\max}^D} \times 10$ .

Así la puntuación del sujeto se recoge totalmente desglosada en los componentes que la han formado, pudiéndose analizar con detalle como se ha llegado al resultado.

### C. Banco de ítems.

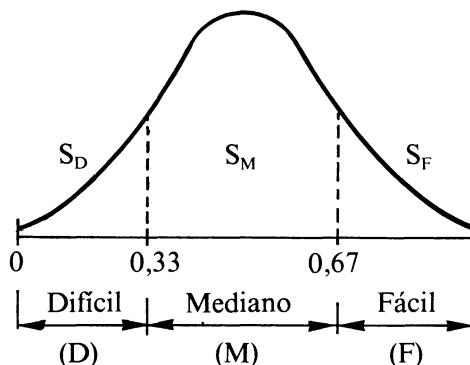
Cada ítem debe ir asociado a una especificación operativa y sólo a una. En consecuencia, queda asociado también a un objetivo específico.

Un ítem sólo se incorpora al banco cuando haya sido testado por primera vez, recogiéndose una serie de datos estadísticos asociados al mismo. De ellos es fundamental la «facilidad», la cual va actualizándose en las sucesivas pruebas en que interviene el ítem, tanto en media ( $f$ ) como en desviación típica ( $\sigma_f$ ).

Para el modelo, más que el valor numérico, como se verá, es básico saber si un ítem es «fácil», «difícil», o «mediano». Por ello se define el concepto «facilidad simbólica» ( $\varphi$ ), que es la proyección de la historia de la facilidad del ítem (supuesta normal) sobre un escalado de la facilidad acotado como indica el gráfico adjunto. Las áreas en que queda cortada la curva normal marcan unos «componentes de la facilidad». Así:

$$\varphi = \frac{S_D}{S_D + S_M + S_F} D + \frac{S_M}{S_D + S_M + S_F} M + \frac{S_F}{S_D + S_M + S_F} F. \text{ Si se tiene, por ejemplo:}$$

$\varphi = 0,230 + 0,73 M + 0,04 F$ , se interpretará como que el ítem es considerado un 23 % difícil, un 73 % mediano y un 4 % fácil. Evidentemente, un ítem recién incorporado al banco, por carecer de  $\sigma_f$  sólo será  $\varphi = D$ ,  $\varphi = M$  o  $\varphi = F$ , sin componentes.



Cuando se elabora un ítem de nueva creación se estima una facilidad simbólica sin componentes igual a la de la especificación operativa.

El archivo se compone así de los siguientes datos:

- Número de ítem
- Descripción
- Alternativa correcta
- Número de especificación operativa asociada
- Facilidad simbólica
- Etc.

#### D. Las pruebas

Es importante la definición de prueba en el modelo F.C.O. «*Una prueba es un conjunto de objetivos que se testan simultáneamente*». Lo anterior se recalca pues en el proceso de corrección de una prueba, el modelo puntuá cada objetivo aisladamente integrando posteriormente los resultados. En síntesis, es cada objetivo individual el que se evalúa, independientemente de cuáles otros estén presentes simultáneamente.

Dado que cada objetivo tiene un peso en la disciplina debería verificarse la condición de que *todos los objetivos tendrían que ser testados a lo largo del curso en pruebas sucesivas*, simplemente por una exigencia aritmética.

El modelo permite, sin embargo, una sutileza en la afirmación anterior transformándola en la de que *todos los objetivos han de ser presentados a test a lo largo del curso*. Entonces, si en una prueba están presentados  $n_1$  objetivos y, en realidad, sólo se testan  $n_2$  ( $n_2 < n_1$ ) existe la técnica de «concentración de pesos» que incrementa los pesos de los objetivos testados, absorbiendo los de los presentados y no testados por reparto proporcional. Se permite así que un objetivo no sea testado si las necesidades del curso lo requieren. Sin embargo, dado que la puntuación final de la disciplina se obtiene por la agregación de las puntuaciones alcanzadas en cada objetivo por sus pesos respectivos, *un objetivo sólo puede ser testado una vez a lo largo del curso*.

Todo lo anterior es válido para las pruebas que denominamos «sumativas», pero el modelo contempla otros tipos de pruebas. De ellas, las más importantes son las pruebas «formativas», las cuales no cumplen las condiciones anteriores pues un mismo objetivo puede ser testado en una formativa y en una sumativa. Sus características fundamentales son que tiene un contenido más reducido y que las puntuaciones de sus objetivos, aunque se obtienen por los mismos algoritmos que las sumativas, sólo se integran en la puntuación total de la disciplina de forma transitoria, considerándose estimaciones de las puntuaciones que obtendría el sujeto en una prueba sumativa de los objetivos involucrados.

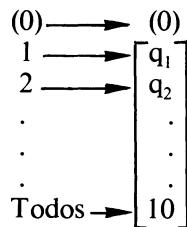
Debido a este carácter de transitoriedad las puntuaciones de los objetivos son sustituidas por las alcanzadas en una prueba sumativa posterior que los teste

nuevamente. El modelo analiza mediante unos índices de control que el nivel de puntuaciones de cada objetivo se eleva en la prueba sumativa respecto a la formativa. Ello indica así que el aprovechamiento del curso es adecuado.

Antes de pasar al contenido de los archivos de pruebas debe comentarse que el modelo posee *tres subsistemas diferentes* relativos a las mismas.

1. *Confección de la prueba.* A partir de los objetivos presentados a test, se determina el número mínimo de especificaciones operativas necesario para alcanzar la representatividad y a continuación se selecciona un ítem (del banco o de nueva creación) para cada una de ellas.

Es de destacar que la prueba queda totalmente definida una vez seleccionados los ítems de cada objetivo. El baremo de puntuaciones por objetivo es el denominado *vector cut-off* el cual especifica la puntuación alcanzada, en función del número de aciertos conseguido. Ello se efectúa posicionalmente. Así, el primer elemento señala la puntuación relativa para un acierto, el segundo la conseguida para dos aciertos, etc. El último elemento del vector (correspondiente a todos los ítems acertados) siempre es 10. Existe, finalmente, una línea virtual que es 0 aciertos → de puntuación relativa, la cual ya no se pone por existir siempre.



El interés del método estriba en que *conocidas las facilidades simbólicas «φ» de los ítems que componen un objetivo*, existen unos modelos matemáticos basados en curvas exponenciales, lineales y logarítmicas que *determinan automáticamente* el vector cut-off correspondiente.

Por otra parte, se determinan automáticamente unos coeficientes denominados *influencias* que realizan las puntuaciones alcanzadas en el caso que el nivel de superación de los objetivos considerados esenciales no haya superado unos mínimos preestablecidos. (Estos sólo son de aplicación en las pruebas sumativas).

Finalmente, el educador, si lo considera conveniente, construye *permutaciones* debiendo establecer tablas de correspondencia para los números de ítem entre una permutación y otra.

2. *Recogida de respuestas.* Las respuestas de los sujetos se introducen al sistema para su evaluación posterior.

3. *Evaluación.* A partir de las características de la prueba (alternativas correctas, permutaciones, vectores cut-off, coeficientes de influencia y pesos por objetivo) se evalua cada objetivo aisladamente, integrándose los resultados posteriormente.

Son de destacar dos puntos de interés:

– Se efectúa siempre corrección por azar, pero no al sistema clásico, sino uno más sofisticado que tiene en cuenta el número de errores producidos sólo en los objetivos no esenciales y efectuándose un reparto proporcional al número de ítems en estos objetivos.

– El número de aciertos (corregido al azar) final se supone que es una estimación de la puntuación verdadera que obtendría el sujeto, denominándosele, «puntuación observada».

En consecuencia, se pueden calcular unas probabilidades, para cada objetivo, de puntuaciones verdaderas de 0 aciertos, 1 acierto, 2 aciertos,... K. Ello se efectúa mediante una aproximación binomial con media la puntuación observada ( $\varphi$ ).

$$p(k) = \binom{n}{k} \varphi^k (1-\varphi)^{n-k}$$

Así se obtiene un vector de probabilidades de número de aciertos. El producto escalar de dicho vector por el vector cut-off da las puntuaciones relativas de los objetivos.

$$\begin{vmatrix} p(1) \\ p(2) \\ \cdot \\ \cdot \\ p(n) \end{vmatrix} \times \begin{vmatrix} q_1 \\ q_2 \\ \cdot \\ \cdot \\ 10 \end{vmatrix} = p(1) \cdot q_1 + p(2) \cdot q_2 + \dots + p(n) \cdot 10 = q$$

En conclusión de todo lo anterior, los archivos de pruebas contienen las siguientes informaciones:

- Tipo de prueba
- Ítems componentes (estos ya marcan las alternativas correctas, los vectores cut-off y los coeficientes de influencia).
- Número de permutación
- Tabla de correspondencia de ítems.

y finalmente:

- Número de sujeto (en el archivo de alumnos)
- Respuestas efectuadas.

## 2. Síntesis de funcionamiento.

Conocidos todos los elementos y archivos componentes del modelo podemos sintetizar el funcionamiento del mismo en los siguientes pasos:

1. Creación del archivo de disciplina.
2. Creación del archivo de alumnos.
3. Creación del banco de ítems.
4. Confección de una prueba.
5. Recogida de respuestas de los sujetos.
6. Evaluación.
7. A partir de los resultados de la evaluación:
  - Confección de gráficos de seguimiento (ver anexo).
  - Actualización de puntuaciones en el archivo de ALUMNOS.
  - Actualización del banco de ítems en cuanto a facilidades y otros datos históricos se refiere.
  - Actualización del archivo de disciplina recogiendo parámetros de control para verificar la bondad de la estructuración del dominio.
  - Resultados auxiliares:
    - Evaluación normativa complementaria.
    - Análisis de normalidad de distribuciones.
    - Valoración de instrumentos.
    - Comparación entre puntuaciones sumativas y formativas, etc.

Como puede observarse, los archivos de disciplina, alumnos y banco de ítems recogen toda la información de las sucesivas pruebas, por lo que éstas, una vez evaluadas, pueden destruirse por carecer de interés posterior.

El modelo ha sido implantado en un ordenador personal Comodore-64 requiriéndose una impresora auxiliar para su funcionamiento.

SUJETO NO. 2

ω 1 [REDACTED]  $7.6 \times 3.39 = ( 25.77 )$   
 ω 2 [REDACTED]  $7.85 \times 6.38 = ( 50.67 )$   
 \*ω 3 [REDACTED]  $4.24 \times 5.46 = ( 23.14 )$   
 \*ω 4 [REDACTED]  $2.66 \times 2.78 = ( 7.38 )$   
 ω 5 [REDACTED]  $3.78 \times 4.14 = ( 15.66 )$

$$\begin{aligned} \text{SUMA PESOS} &= 22.15 \\ (\text{TOTAL}) &= ( 122.02 ) \\ \text{BASE 10} &= 5.51 \end{aligned}$$

SUJETO NO. 5

ω 1 [REDACTED]  $5.68 \times 3.39 = ( 19.26 )$   
 ω 2 [REDACTED]  $3.33 \times 6.38 = ( 21.22 )$   
 \*ω 3 [REDACTED]  $5.19 \times 5.46 = ( 28.32 )$   
 \*ω 4 [REDACTED]  $5.03 \times 2.78 = ( 13.98 )$   
 ω 5 [REDACTED]  $6.98 \times 4.14 = ( 28.88 )$

$$\begin{aligned} \text{SUMA PESOS} &= 22.15 \\ (\text{TOTAL}) &= ( 111.66 ) \\ \text{BASE 10} &= 5.04 \end{aligned}$$

SUJETO NO. 11

ω 1 [REDACTED]  $2.84 \times 3.39 = ( 9.62 )$   
 ω 2 [REDACTED]  $3.03 \times 6.38 = ( 19.33 )$   
 \*ω 3 [REDACTED]  $3.35 \times 5.46 = ( 18.29 )$   
 \*ω 4 [REDACTED]  $1.39 \times 2.78 = ( 3.87 )$   
 ω 5 [REDACTED]  $1.01 \times 4.14 = ( 4.16 )$

■■■ .77 ( 2.62 )  
 ■■■ .82 ( 5.26 )  
 ■■■ .27 ( 1.13 )  
 ■■■ ( 9.01 )  
 ■■■ .41

$$\begin{aligned} \text{SUMA PESOS} &= 22.15 \\ (\text{TOTAL}) &= ( 55.26 ) \\ \text{BASE 10} &= 2.49 \end{aligned}$$

SUJETO NO. 20

ω 1 [REDACTED]  $4.48 \times 3.39 = ( 15.17 )$   
 ω 2 [REDACTED]  $7.5 \times 6.38 = ( 47.85 )$   
 \*ω 3 [REDACTED]  $8.6 \times 5.46 = ( 46.96 )$   
 \*ω 4 [REDACTED]  $10. \times 2.78 = ( 27.8 )$   
 ω 5 [REDACTED]  $8.15 \times 4.14 = ( 33.74 )$

$$\begin{aligned} \text{SUMA PESOS} &= 22.15 \\ (\text{TOTAL}) &= ( 171.53 ) \\ \text{BASE 10} &= 7.74 \end{aligned}$$

## REFERENCIAS BIBLIOGRÁFICAS

- BERK, R.A. (1980): A Framework for Methodological advances in Criterion Referenced Testing. *Applied Psychological Measurement*, 1, 4, 563-573.
- BLOCK, J.H. (1971): Criterion referenced measurement: Potential. *School Review*, 69, 289-298.
- BRENNAN & KANE. (1977): An Index of Dependability for Mastery tests. *Journal of Education Measurement*, 1977, 14, (3), 277-289.
- CRONBACH, L.J. (1970): Review of Bormuth's on the theory of Achievement test Items. *Psychometrika*, 35, 509-511.
- CRONBACH; GLASER; NANDA Y RAJARATHAM. (1972): *The dependability of Behavioral Measurement*. John Wiley & Sons, New York.
- GLASER, R. (1963): Instructional Technology and the measurement of learning outcomes. *American Psychologist*, 18, 519-521.
- GLASER, R., y NITKO, A.J. «Measurement i learning and instruction» In R.L. Thorndike (Ed.) (1971): *Educational measurement*. (2nd. ed.), Washington, American Council on Education.
- GLASS, G. (1978): Standars and criteria. *Journal of Educational Measurement*, 13, 4, 237-261.
- GRAY, (1978): A comparasion of Piagetian theory and criterion-referenced measurement *Review of Educational Research*, 48, 223-249.
- HAMBLETON, R.K. y NOVICK, M.R. (1973): Toward and integration of Theory and Method for Criterion-referenced Tests. *Journal of Educational Measurement*, 10, 159-170.
- HAMBLETON; SWAMINATHAM; ALGINA Y COULSON. (1978): Criterion-referenced testing an measurement: A review of technical issues and developments. *Review of Educational Research*, 48, 1-47.
- HAMBLETON, R.K. (1980): Contributions to Criterion-Referenced Testing Technology: An Instro-dunction: *Applied Psychological Measurement*, 4, 4, 421-424.
- HIVELY, W.; PATTERSON, H.L. Y PAGE, S.A. (1968) «A «universe-defined» System of arithmetic achievement test. *Journal of Educational Measurement*, 5, 275-290.
- HUYNH, H. (1976): On consistency of Decision in Criterion-Referenced Testing. *Journal of Educational Measurement*, 13, 253-264.
- KANE, M.T. y BRENNAN, R.L. (1980): Agreement Coefficients as indices of Dependability for Domain-References Test. *Applied Psychological Measurement*, 4, 1, 105-126.
- LINN, R. (1980): Issues of Validity of Criterion-Referenced Measures. *Applied Psychological Measurement*, 4, 4, 574-561.
- LIVINGSTON, S.A. (1972): Criterion-Referenced Applications of Classical test Theory. *Journal of Educational Measurement*, 9, 13-26.
- LORD, F.M. y NOVICK, M.R. (1968): *Statistical theories of mental test Scores*. Reading M.A. Addison, Wesley.

- MARTÍNEZ, M.R. (1981): «Principios Psicométricos de la técnica en evaluación conductual» En Fernández Ballesteros, R. y Carrobles, J.A. (Eds.), *Evaluación conductual. Metodología y Aplicaciones*. Pirámide, Madrid, 157-198.
- MEHERENS, W. y LEHMANN. (1979): *Measurement and evaluation in Education and Psychology*, 1973, 1978. Traducida al castellano por Compañía Editorial Continental, S.A. México, *Medición y Evaluación en la Educación y en la Psicología*, 1982.
- MESKAUSKAS, J.A. (1976): Evaluation Models of Criterion-referenced testing: Views regarding Mastery and Standard-Setting. *Review of Educational Research*, 46, 1. 133-158.
- MILLMAN, J. (1974): «Criterion-referenced measurement» On J. Popham, W.J. (Ed.), *Evaluation in education: Current applications*. Berkeley, California: Mc. Cutcham Publishing Co..
- NITKO, A.J. (1980): Distinguishing the many varieties of Criterion-Referenced tests. *Review of Educational Research*, 50, 461-485.
- POPHAM, W.J. y HUSEK, T.R. (1969): Implications of Criterion-Referenced Measurement. *Journal of Educational Measurement*, 6, 1-9.
- POPHAM, W.J. (1983): *Evaluación basado en criterios*. Magistrado Español, Madrid.
- SHEPARD, L. (1980): Standard Setting Issues and Methods. *Applied Psychological Measurement*, 4, 447-467.
- SUBKOVIACK, M.J. (1976): Estimating reability from a single administration of a mastery test. *Journal of Educational Measurement*, 13, 265-276.
- SWAMINATHAN, H.; HAMBLETON, R.K. y ALGINA. (1974): Reability of Criterion-Referenced tests: A decision-theoretic formulation. *Journal of educational Measurement*, 11, 263-276.
- TENBRINK, T.D. (1981): *Evaluación. Guía práctica para profesores*. Narcea, Madrid. En el original en inglés, *Evaluation: A practical guide for teachers*.
- TRAUB, R.E. y ROWLEY, G.L. (1980): Reability of test Scores and Decisions. *Applied Psychological Measurement*, 4, 517-545.

## Problemática de la enseñanza universitaria

por  
Fco Javier Tejedor

### LA EVALUACIÓN EN LA UNIVERSIDAD.

En cualquier nivel educativo el análisis de un «plan evaluativo» debería contemplar, al menos, referencias a los alumnos, a los programas y a los profesores. Sin embargo, creo, hay una cierta tendencia en este país a limitar nuestro estudio al análisis del rendimiento de los alumnos (cuando escribo estas líneas no conozco el contenido de los documentos que nos van a presentar a discusión y me temo que la mayoría de ellos estén pensados en relación con el análisis del rendimiento de los alumnos).

No quiero decir que no sea un tema importante a considerar. Simplemente me interesa poner de manifiesto que no es el único aspecto a considerar, por varias razones:

- En primer lugar, porque no creo que pueda llevarse a cabo con eficacia la evaluación aislada, independiente, de los procesos de aprendizaje desconectados de los procesos de enseñanza (Gimeno, J. 1981).
- En segundo lugar, porque no comparto la idea (creo que poco la mantienen todavía) de pensar que residen en el alumno todas las potencialidades de logros a conseguir y todas las responsabilidades por logros no conseguidos.
- En tercer lugar, porque no es fácil establecer interpretaciones recursivas; me explico: una vez emitido el juicio, la valoración sobre el rendimiento a partir de procedimientos objetivos o subjetivos, vinculados o no a procesos de estandarización, expresados cualitativa o cuantitativamente,..., no resulta fácil, porque equivaldría a volver al principio, considerar si los objetivos, conseguidos o no, que determinaron esa valoración estaban suficientemente explicitados, adecuadamente formulados,...; si las actividades desarrolladas fueron las más adecuadas en función de los recursos disponibles, de los intereses, capacidades y motivaciones de los alum-

nos,...; si las pruebas de evaluación eran adecuadas,..; si el profesor está suficientemente capacitado para ejercer la función docente y, aún así, si «puso de su parte» todo lo que podía,...

Esto no significa, claro, que cualquier estudio sobre la evaluación del rendimiento tenga necesariamente que considerar al mismo tiempo todos los aspectos implicados en los procesos de enseñanza-aprendizaje. Puedo aceptar sin traumas la posibilidad de analizar los componentes de un sistema como método de aproximación para su comprensión y explicación; es una actitud posibilista, analítica, pero no simplificadora.

Retomemos el punto de partida con referencias exclusivas ya al ámbito universitario.

No parece muy arriesgado afirmar que la evaluación en nuestra universidad se ha referido casi exclusivamente a la evaluación del rendimiento de los alumnos. Aunque puede pensarse en características diferenciales específicas de los alumnos universitarios y, por supuesto, en prácticas evaluativas también diferenciales, aunque menos específicas, no creo que ello deba suponer modelos de investigación del rendimiento diferentes a los de otros niveles educativos (tengo mis dudas al hacer esta afirmación; me gustaría conocer otras opiniones sobre todo aquellas que no sean coincidentes).

Podrá cambiar la naturaleza y el número de variables a considerar; su ponderación u orden de entrada en el modelo. Podrán, y deberán, cambiar los instrumentos de medición... Pero los modelos y las técnicas estadísticas consiguientes, entiendo, no tienen porqué cambiar. Entre otras razones por que no hay otros. Si acaso yo diría que su aplicación en el ámbito universitario se simplificada limitación, casi exclusivamente sumativa, de la evaluación que se realiza.

Como modelos ya se han presentado en este Seminario («Modelos de investigación sobre el rendimiento. Problemática y tendencias actuales») les remito al trabajo de mi compañero Rodríguez Espinar.

Respecto a la evaluación de programas puede decirse otro tanto. Aunque no aparece tan explicitado como en el caso anterior, entiendo que esta problemática se incluye en el tópico «Rendimiento y análisis de Sistemas Educativos».

El objetivo básico de la evaluación de programas parece claro: averiguar si a través de un programa se alcanza lo que se intenta. Lo cual no es fácil de comprobar pues a menudo ocurre que los fines a conseguir no se han explicitado, son incoherentes o son tan limitados que su consecución no significa casi nada. Podría ocurrir también que estando bien determinados los fines, no se implementen los procedimientos para su consecución.

La evaluación de programas debe llevarse a cabo de un modo sistemático. Para ello se han elaborado numerosos modelos a lo largo de las últimas décadas (Tyler, R. 1942; Stake, R.E. 1967; Scriven M. 1972; Weis, C.A. 1972; Wortman, P.M. 1975). Las coincidencias entre ellos parecen reducirse a entender el proceso como una serie de actividades de retroalimentación interconexas que ocurren en función de la interacción de tres componentes: unidades de organización, conceptos teóricos fundamentales y procesos evaluativos.

Landsheere, G. (1982) expone y comenta con detalle el modelo de Stake, un modelo cualitativo de evaluación de programas que denomina «evaluación respondiente» y que viene, en su opinión, a subsanar las deficiencias de la investigación nomotética en este campo.

El modelo mas empleado, quizá, en el ámbito universitario ha sido el modelo CIPP –contexto, input, proceso, producto– (Stufflebeam y otros, 1971):

- a) al evaluar el contexto pretende establecerse si los objetivos propuestos son adecuados.
- b) al evaluar el input, si el programa se adecua a las necesidades que quiere satisfacer.
- c) al evaluar el proceso, si la puesta en práctica es correcta
- d) al evaluar el producto, si el programa «funciona».

Wergin, J.F. (1977) indica los criterios que conviene tener en cuenta en cada una de las fases: En la fase a), los criterios hacen referencia a las características e intereses de los alumnos a los que el programa se dirige; a la valoración del programa frente a otras alternativas, a los medios y recursos,... En la fase b) los criterios hacen referencia a si las metas son explícitas y consegibles y a si las actividades son congruentes con las metas y posibilitan su consecución. En la fase c), los criterios señalados son el de la instrumentalización (intenta averiguar si la puesta en marcha del programa se realiza siguiendo las pautas que han sido propuestas y cuales son los obstáculos aparecidos) y el de la reacción de las partes implicadas en el proceso, profesores y alumnos. En la fase d) se señalan los criterios para contrastar la eficacia del programa, tanto en sus efectos directos como indirectos.

El problema de la evaluación de profesores, que analizaré con mas detalle, creo que adquiere una relevancia especial en el contexto universitario, tanto por el decisivo papel que el profesor puede desempeñar como por entender que tomando como punto de referencia la eficacia del profesor pueden detectarse las ventajas e inconvenientes de un determinado programa, la bondad con que se pone en práctica y el progreso logrado por los alumnos.

## **LA EVALUACIÓN DE LA EFICACIA DOCENTE Y/O EVALUACIÓN DEL PROFESOR.**

La vertiente metodológica de este problema ya fue tratada en un excelente trabajo presentado en el Congreso Nacional de Granada (Escudero, J.M. 1980).

Creo que la persistencia histórica en nuestro país del sistema de oposiciones para acceder, con carácter vitalicio, a la situación docente ha tenido y tiene mucho que ver con la despreocupación casi absoluta por las prácticas de evaluación de la docencia universitaria, ya que ello implicaba inexcusablemente la evaluación del profesor. Al imposibilitar el sistema de oposiciones la viabilidad de una

de las dos consecuencias fundamentales que deben derivarse de un proceso evaluativo en la docencia universitaria, a saber, la promoción del profesorado, la otra, la puesta en marcha de un mecanismo formativo de retroalimentación que ayude al profesor a mejorar su actuación decente, se ha desconsiderado de facto, aunque no se explice, por irrelevante.

En la mayoría de los países ha sido, creo, la necesidad de ampliar las plantillas docentes y la necesidad permanente de «rendir cuentas» lo que ha puesto en marcha los procesos de evaluación de la eficacia docente y con ellos, claro, del profesor.

El presente nos muestra algunos indicios de que la situación española puede cambiar; si bien persisten las oposiciones (aunque mucho más «axeitadas»), la totalidad de los estatutos universitarios de reciente implantación contemplan la obligatoriedad de llevar a cabo evaluaciones del profesorado, que deberán contemplar referencias a las obligaciones mínimas (elaboración del programa, asistencia a clase, impartición del programa, atención a los alumnos, asistencia a las reuniones del departamento, cumplimiento de las dedicación,...) y a las valoraciones de la eficacia docente. Los resultados de esta evaluación deberán obligatoriamente hacerse llegar a los tribunales que juzguen los concursos de acceso a las plazas docentes universitarias (artículo 8, punto 4, del Decreto 1888/84 de 26 de septiembre, BOE de 26-X-84).

Pese a que la evaluación de la enseñanza universitaria se viene realizando hace años en numerosos países es todavía un problema con importantes limitaciones, tanto teóricas como prácticas. Limitaciones teóricas porque no hay acuerdo unánime sobre las finalidades de la enseñanza universitaria (lo que tiene relación con el problema que hoy «flota» en el contexto social y político respecto a la alternativa entre la «universidad técnica o profesional» y la «universidad cultural), ni se dispone de un modelo ideal de profesor universalmente aceptado que nos pueda servir como marco de referencia para un estudio evaluativo.

Limitaciones prácticas porque es difícil elegir la estrategia evaluativa adecuada ya que su validez ha de establecerse indirectamente ante la mencionada carencia de un modelo teórico que explice con suficiente claridad la relación que se establece entre las distintas variables contempladas en el modelo.

Aunque el tema de los modelos de profesor ya lo tratamos en un escrito anterior sobre la evaluación del profesorado (Aparicio, J., Sanmartín, R. y Tejedor, F.J. 1982; Gimeno Sacristán, J. 1983), nos parece oportuno incluir aquí algunas precisiones.

Pueden distinguirse tres paradigmas que sustentan las investigaciones sobre lo que debe ser un profesor efectivo (Doyle, W. 1977) que a su vez inspiran las diferentes estrategias de evaluación de la enseñanza: el paradigma proceso-producto, el paradigma mediacional y el paradigma ecológico.

En el paradigma proceso-producto tratan de establecerse relaciones entre el proceso (conductas del profesor) y el producto (resultados medibles que alcanzan los alumnos). El conjunto de conductas del profesor que estén más relacio-

nadas con los resultados deseables obtenidos por el alumno definirían al profesor efectivo. Este paradigma, derivado de las corrientes conductistas en psicología, es un modelo explicativo muy limitado: no proporciona las claves para averiguar por qué ciertas conductas del profesor están relacionadas con ciertos resultados alcanzados por los alumnos y, lo que es peor, si se producen discrepancias entre los resultados de las diferentes investigaciones no habrá modo de interpretarlas. Como de hecho sucede que unos estudios muestran que determinadas conductas del profesor están relacionadas con un determinado efecto, mientras que en otros se demuestra que son otras las conductas del profesor las relacionadas con ese mismo efecto (McNeil, J.D. y Popham, W.J. 1973; Rosenshine, B. y Furst, N. 1971), este paradigma parece insuficiente.

Las limitaciones del paradigma proceso-producto derivan de pensar que lo haga el profesor es la causa inmediata de los resultados obtenidos por el alumno. Los investigadores inspirados en la psicología cognitiva han incorporado la idea de que lo que el profesor hace no influye de un modo directo en los resultados alcanzados por los alumnos, sino que más bien influyen en ciertos procesos que constituyen la actividad cognitiva de los alumnos. Dicha actividad cognitiva da lugar a su vez a los resultados logrados por los estudiantes. De este enfoque se deriva un segundo paradigma en las investigaciones sobre la efectividad del profesor que suele denominarse paradigma de los procesos mediadores y que se diferencia del anterior en que introduce un tercer término en la relación profesor-alumno que corresponde a las actividades cognitivas a las que se entrega el alumno cuando procesa la información que se le presenta.

El paradigma de los procesos mediadores trata de aplicar a la realidad de la enseñanza las conclusiones de los trabajos realizados en los laboratorios de psicología sobre conocimiento y memoria. La generalización de los resultados del laboratorio (obtenidos con rigurosas pautas experimentales) a la complejidad real del contexto educativo es la limitación más importante de este paradigma.

Limitación que pretende superar el tercero de los paradigmas al que hemos hecho referencia: el paradigma ecológico. Se trata ahora de establecer la eficacia de un determinado profesor con respecto a unos determinados alumnos en un marco de clase específico. Los investigadores que siguen este paradigma suelen utilizar estrategias naturalistas, etnográficas, con objetivos fundamentalmente comprensivos y descriptivos.

De este enfoque ecológico no puede derivarse un modelo de profesor con aplicación a una amplia gama de situaciones de enseñanza. Por eso, quizás, convenga entenderlo como complementario a otros enfoques, sobre todo por ampliar las competencias que se entiende debería alcanzar el alumno.

No parece haber ninguna duda de que todos ellos tienen ventajas e inconvenientes. La mejor solución pasa por integrarlos adecuadamente cuando eso sea posible. Esta es la línea que entiendo sugiere Foster, D.C. (1984) al presentar 199 indicadores para una enseñanza eficaz y posteriormente reducirlos a 47 agrupados en tres categorías: planificación, instrucción y responsabilidad. La evaluación de estos indicadores puede llevarse a cabo siguiendo diferentes estra-

tegias, aunque como señalamos anteriormente no es fácil superar las dificultades intrínsecas para su integración, pues se sigue considerando el paradigma proceso-producto como medida más directa de la competencia docente, por pretender básicamente relacionar los logros estudiantiles con determinadas conductas del profesor.

### **El rendimiento de los alumnos como medida de la competencia docente.**

La procedencia de considerar el rendimiento de los alumnos como media de la competencia docente es un tema muy polémico. Si por una parte no parece haber ninguna duda de que el objetivo fundamental de la tarea docente es facilitar el aprendizaje de los alumnos, resulta lógico pensar que una buena manera de saber si un profesor es efectivo es averiguar lo que sus alumnos han aprendido. Por otra, se cuestiona cómo son y qué miden los exámenes: si responden a una valoración del aprendizaje a corto o largo plazo, si son adecuados con relación a los objetivos fijados... Recomiendo la lectura del artículo de Millman, J. (1981) así como la de las referencias que allí se incluyen.

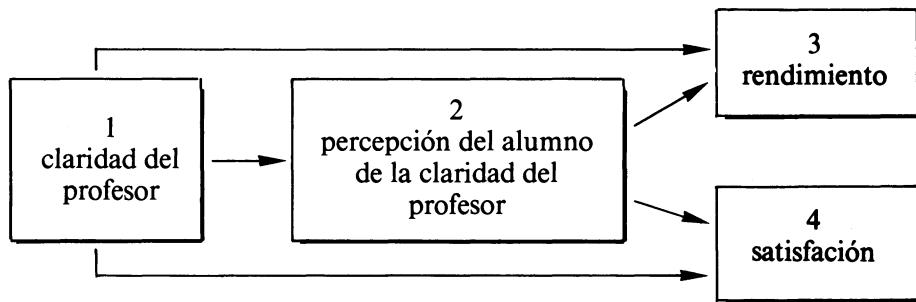
Lo que si está claro es que existen procedimientos para convertir el rendimiento estudiantil en valoraciones acerca de la efectividad del profesor (el rendimiento estudiantil como variable explicativa «*a posteriori*» de eficacia docente); procedimientos que responden a diversas situaciones de investigación y que utilizan técnicas estadísticas también diversas (básicamente, el análisis de covarianza y el análisis de regresión múltiple).

Son numerosos los estudios que relacionan determinadas destrezas o capacidades del profesor con el rendimiento estudiantil, manteniendo estas variables un status más «ortodoxo»; las destrezas o capacidades del profesor como variables independientes o predictoras y el rendimiento estudiantil como variable dependiente o criterio. Citaremos algunos estudios recientes, sin pretender ser exhaustivo pues esta vertiente ha sido también considerada en este Seminario (La investigación sobre el rendimiento: Síntesis informativa y estudios de metaanálisis).

El trabajo de Theodory, G.C. y Day, R.C. (1985), «Relación del estilo docente, rasgos de ansiedad y experiencia del profesor con el nivel alcanzado por los alumnos», pone de manifiesto la existencia de una correlación negativa significativa entre la ansiedad del profesor y el nivel alcanzado por los alumnos así como una correlación positiva entre la experiencia del profesor y el nivel alcanzado por los alumnos; en ambos casos, con la variable estilo docente como moduladora ya que esas relaciones solo aparecen cuando es el estilo responde a una disposición favorable. Al aplicar el análisis de varianza, se encuentra la ansiedad como única variable de influencia significativa.

El trabajo de Hines, C. y otros (1985), «Claridad del profesor y sus relaciones con el rendimiento y la satisfacción de los estudiantes», donde utilizan la observación y el cuestionario como método para la obtención de datos y la correla-

ción canónica, la regresión múltiple y el path análisis como técnicas estadísticas adecuadas para su análisis. Resulta muy curiosa la observación del diagrama path que presentan:



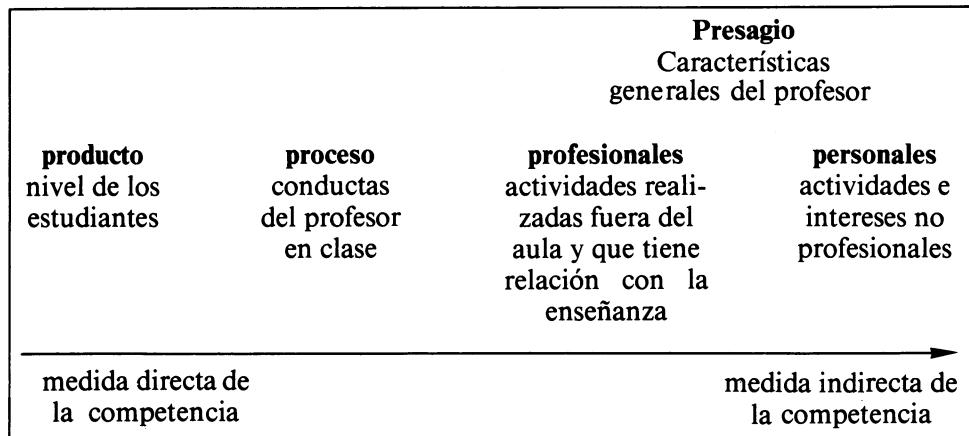
No hay efecto directo entre las variables 1 y 4 ( $-0.05$ ) pero si un importante efecto indirecto a través de la variable 2 (0,51), lo que en conjunto da cuenta de la correlación empírica entre las variables 1 y 4 (0,46).

Otros trabajos que siguen igualmente esta línea metodológica serían: «Utilización del lenguaje audiovisual en la enseñanza de la estadística» (Bartolomé, A. y Mateo, J. 1983), que sigue un diseño experimental clásico con grupo experimental, grupo placebo y grupo control; «Aptitudes mentales y rendimiento universitario» (Chorro, J.L. 1977); «Factores del rendimiento académico en la Universidad» (Touron, J. 1984); «Factores y causas del fracaso académico en la Universidad» (Marín, M. 1983). En algunos casos (Nieto, J.M. 1982) se señala el rendimiento del alumno como variable dependiente y se señalan dos tipos de variables independientes relacionadas con el profesor: su ser (variables estructurales) y su actuar (variables dinámicas). A veces no se estudia el rendimiento directamente sino determinadas situaciones que lo explican o justifican (Bean, J.P. 1985); «Modelo causal explicativo del síndrome de abatimiento». Como variables exógenas se incluyen los factores académicos, psicológicos y ambientales; como variable endógena la socialización en clase; como variable criterio, el abatimiento.

En otros casos (Rivas, F. y Monera, M.L. 1984) se buscan relaciones entre el rendimiento de los alumnos universitarios y variables más antecedentes que la actuación del profesor: «Rendimiento en la universidad en función del seguimiento o no del consejo vocacional recibido en COU». Trabajo muy interesante tanto por su temática como por la metodología utilizada (el análisis discriminante). Los resultados que se obtienen son muy sugerentes.

Me gustaría que respecto al tema «competencia docente/rendimiento de los alumnos» al menos quede claro un principio fundamental: la evaluación de profesores será deficiente si se usa como único criterio el rendimiento final de los

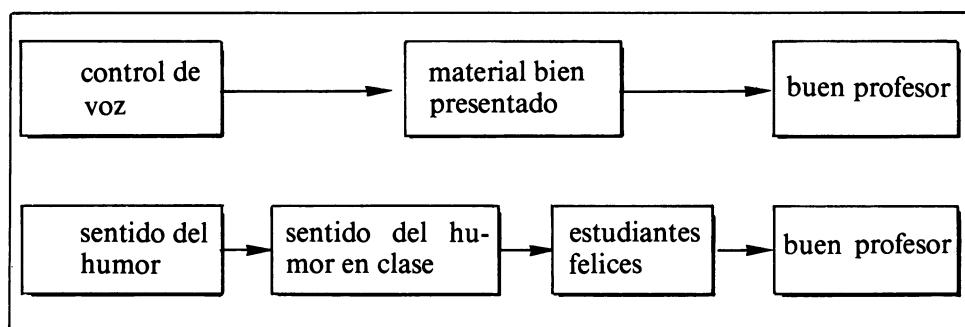
estudiantes. Entiendo como necesario la ampliación de los campos de referencia; si se prefiere, como algunos opinan, entiendo como necesario la obtención de «medidas indirectas» de la competencia docente, generando, lo que King, J.A. (1981) denomina un «modelo multidimensional de evaluación» al tiempo que sugiere la relación entre los componentes del modelo de evaluación de la competencia docente y el tipo de medida que implica:



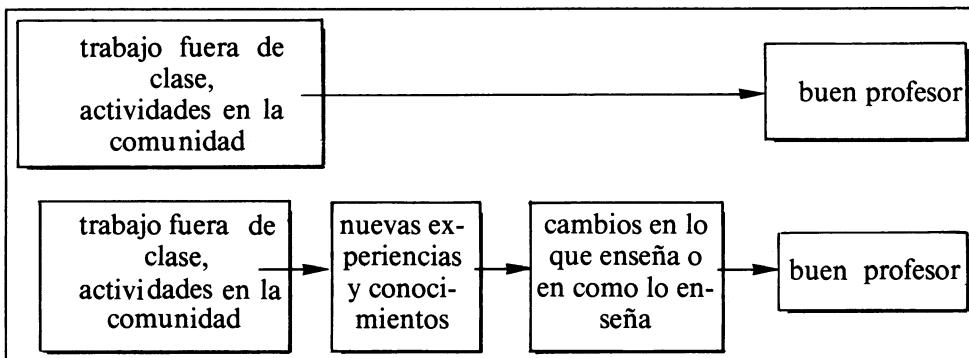
(King J.M. obra citada, p. 169).

Las medidas indirectas, sigue diciendo King, deben entenderse como complementarias, a fin de no caer en «conclusiones fáciles» de las que nos indica algunos ejemplos:

a) En variables presagio:



b) En variables profesionales o personales:



#### Otros métodos para la evaluación de profesores en la enseñanza universitaria.

Es muy variada la gama de métodos que se han venido utilizando en la evaluación de profesores en la enseñanza universitaria: desde los registros en vídeo hasta las valoraciones estudiantiles pasando por los auto-informes, las pruebas estandarizadas o los juicios de los colegas.

La grabación en vídeo de la actuación de un profesor, al permitir llevar a cabo una observación y sistemática del comportamiento docente en el aula, posibilitando la obtención de registros fiables y válidos (sobre todo cuando se supera el impacto de la presencia de la cámara), es práctica muy frecuente. Algunos autores (Centra, J.A. 1977) opinan que es el mejor método para conseguir mejorar la labor del profesor siempre que las grabaciones sean discutidas por el profesor, un especialista y un colega.

El auto-informe consiste en que los profesores redacten un texto sobre su actuación docente. Estas autoevaluaciones de los profesores pueden girar en torno a si los objetivos propuestos, los materiales o los procedimientos empleados han sido adecuados; generalmente incluyen datos de interés como las actividades de todo tipo llevadas a cabo y el trabajo de investigación realizado.

En contra de lo que cabría esperar, las autoevaluaciones no se caracterizan por su parcialidad en favor del profesor. No pocos estudios muestran que los datos obtenidos con las autoevaluaciones se aproximan bastante a los que resultan aplicando otras básicas (Braskamp, L.A. y otros, 1979; Marsh, H.W. y otros, 1979).

Las pruebas estandarizadas para la evaluación de profesores (testing teachers) se ofrecen como alternativa para la mejora de los resultados obtenidos (Peterson, K. 1984), o como específicamente idóneas para la evaluación de algunas de las dimensiones del comportamiento docente (Pugach, MC. y Raths, J.D. 1983; Sandefur, J.T. 1983).

Las evaluaciones de los colegas, acompañadas o no de observación del comportamiento del profesor en el aula, son especialmente útiles para evaluar algunos aspectos como la labor investigadora, el profesionalismo del profesor o su dominio de la asignatura. En general, los profesores pueden estimar con bastante exactitud la competencia de un colega no tanto en lo que se refiere al modo en que da la clase, sino en lo que respecta al bagaje de conocimientos que posee y a su preparación científica y pedagógica. Braskamp, L.A. (1978) ha llevado a cabo un estudio en el que se pone de manifiesto la ventaja de este método evaluativo cuando se pretende evaluar la capacidad del profesor para organizar un curso, para elegir los materiales adecuados o para dar una orientación correcta desde el punto de vista científico.

Las evaluaciones de los estudiantes, en opinión de algunos autores (Banister, J. y otros, 1961; Rayder, N.F. 1968) proporcionan, en conjunto, el mejor criterio de calidad de la instrucción dado que ellos son los más «fieles observadores». Sin duda, es el método más utilizado (Moomaw, M.T. 1977). Puede llevarse a cabo bien a partir de comentarios escritos, entrevistas de grupo o escalas objetivas de valoración.

Los comentarios escritos ofrecen una importante información diagnóstica (Tuckman, B.Q. y Oliver, W.F. 1968) pero de muy difícil catalogación por el amplio campo que abarcan, distorsionando los análisis posteriores.

Las entrevistas en grupo, dirigidas por un monitor, ofrecen interesantes posibilidades informativas sobre la labor docente. Su registro posibilita la catalogación y el análisis posterior (Ory, J.C. y otros, 1980).

Los cuestionarios o escalas de valoración se configuran con un determinado número de preguntas cerradas en las que se valora al profesor en relación con distintos atributos relevantes desde el punto de vista de la eficacia docente. Se pide a los estudiantes que valoren al profesor en cada uno de los ítems del cuestionario. En un trabajo nuestro ya citado (Aparicio, J. Sanmartín, R. y Tejedor, F. 1982) se tratan con detalle todos los pormenores sobre la problemática de la utilización de este tipo de cuestionarios: selección y número de ítems, número de categorías de cada ítem, dimensiones resultantes, fiabilidad y validez de los cuestionarios, factores que pueden afectar a las valoraciones estudiantiles (tamaño de la clase, tipo de asignatura, momento de aplicación,...), etc. También se incluyen las críticas más frecuentes a su uso.

Sobre esta cuestión, Wittrock, M.C. y Lumsdaine, A.A. (1977) opinan que la mayoría de las investigaciones sobre las valoraciones que los estudiantes hacen del profesor no indican que las valoraciones midan la efectividad del profesor, la calidad de la enseñanza o el rendimiento. Lo que miden estas valoraciones en realidad, en su opinión, es la satisfacción de los estudiantes, las actitudes de éstos hacia sus profesores y sus clases, sus necesidades psicosociales y las características de la personalidad del profesor, su popularidad y sus cualidades como expositor.

A los argumentos de apoyo que expusimos en 1982 habría que añadir algunos más (Aleamoni L.M. 1981) pues sigo pensando que la utilización de este tipo de cuestionarios es del todo procedente:

a) Se dice que los estudiantes no pueden emitir juicios competentes sobre el profesor y la enseñanza por su inmadurez y falta de experiencia. Los datos de la investigación no confirman esta opinión pues en gran número de estudios se encuentran correlaciones de 0,87 a 0,89 en las valoraciones sucesivas (Guthrie, E.R. 1954). El intervalo correlacional obtenido en de 0,70 a 87 en los estudios de Costin, F. y otros (1971), Gillmore, G.M. (1973) y Hogan, T.P. (1973).

b) Se dice que existe una gran influencia sobre la evaluación en base a la fama que el profesor tenga por sus publicaciones. Efectivamente la investigación y la docencia difícilmente se pueden separar: una buena enseñanza y una buena investigación están tan unidas que es innecesario evaluarlas independientemente (Borgatta, F.F. 1970; Deming, W. E. 1972). Los resultados de la investigación son dispares en este punto: mientras que algunas muestran una relación positiva entre productividad investigadora y eficacia docente (Maslow, A.H. y Zimmerman, W. 1956; McGrath, E.J. 1962), McDaniel, E.D. y Feldhusen, J.F. y Feldhusen, J.F. 1970), otras no encuentran relaciones significativas (Voecks, V.W. 1962; Hayes, J.R. 1971; Aleamoni, L.M. y Yimer, M. 1973; Linsky, A.S. y Straus, M.A. (1975). Tampoco aparecen relaciones significativas cuando la valoración es realizada por los colegas en lugar de por los estudiantes (Aleamoni, L.M. y Yimer, M. 1973).

c) Se dice que la popularidad del profesor juega un papel importante en las valoraciones estudiantiles. Efectivamente así ocurre para algunas de las dimensiones del cuestionario, pero se mantiene la capacidad discriminativa de los juicios para otras dimensiones (Grush, J.E. y Costin, F. 1975; Ware, J.E. y Williams, R.G. 1977; Frey, P.W. 1978; Perry, R.P. y otros, 1979).

d) Se dice que los juicios estudiantiles no son exactos hasta que han pasado varios años desde que terminó el curso o la universidad. Los estudios de Drucker, A.J. y Remmer, H.H. (1951) muestran que valoraciones de antiguos alumnos coinciden con la de los alumnos del momento. Estos resultados se confirman en trabajos de Marsh, H.W. (1977), Marsh, H.W. y Overall, J.V. (1979) y McKeachie, W.J. y otros (1978).

e) Se manifiesta la dificultad de incorporar los resultados de la valoración estudiantil a la enseñanza. Creo que puede ocurrir así cuando se entiende la evaluación del profesor como sumativa y sancionadora en lugar de formativa. En cualquier caso, los estudios de metaanálisis de Cohen, P.A. (1982) ponen de manifiesto que las relaciones encontradas entre las valoraciones estudiantiles y la eficacia docente son más fuertes y consistentes de lo que cabía esperar.

Otras críticas, más suaves, pero interesantes por su proximidad y sobre todo porque nacen de la experiencia (Villar, L.M. 1983) han sido puestas de manifiesto por Villar, L.M. (1985) en un excelente trabajo sobre «Evaluación del ambiente de aprendizaje en el aula universitaria». Pero siendo interesantes sus críticas mucho más me parece la alternativa que sugiere para acercarse a la evaluación del profesorado con el fin de modificar conductas docentes y ambientes de enseñanza para, en última instancia, mejorar el rendimiento estudiantil.

Esta nueva alternativa metodológica, la evaluación de ambientes de clase, se configuró adecuadamente a partir de la revisión que Anderson, C.S. publicó en

1982. Incluye estrategias etnográficas (variables relevantes, utilización de observaciones y entrevistas en profundidad,...) concebidas para la obtención de datos sobre el «ambiente de clase» junto a técnicas estadísticas para el análisis de datos y la obtención de las dimensiones subyacentes (análisis factorial, análisis clusters,...).

No desaparecen con ella los problemas que antes hemos mencionado, pero entiendo que se trata de un procedimiento complementario, de una posibilidad mas de obtener información sobre la actuación del profesor universitario en el aula.

De cualquier forma, esta diversidad metodológica en las distintas fases del proceso, con la debida integración entre la mismas, parece ser el mejor camino para evitar que cualquiera de las estrategias utilizada separadamente puede ser juzgada como parcial e irrelevante, al tiempo que se facilita la conexión entre la teoría y la práctica; al tiempo que se posibilita la comprobación empírica del modelo teórico. Si para esta comprobación utilizamos adecuadamente el modelo estadístico que mejor se adapte a las características de los datos (teniendo en cuenta que la calidad de una técnica depende no sólo de sus características inherentes sino también del contexto y manera en que se implemente), estaremos en condiciones de comenzar a sentar las bases sólidas de un conocimiento auténticamente científico sobre la eficacia docente, uno de los problemas cruciales de la problemática educativa (McGreal, T.L. 1982). Esta es la línea que pretendo seguir en un trabajo recién comenzado en la Facultad de Químicas de la Universidad de Santiago.

#### REFERENCIAS BIBLIOGRÁFICAS

- ALEAMONI, L.M. (1981): «Student ratings of instruction» en J. Millman, *Hand-boock of teacher evaluation*, Sage Publications, Beverly Hill, California.
- ALEAMONI, L.M. y YIMER, M. (1978); «An investigation of the relationship between colleague rating, student rating, research productivity, and academic rank in rating instructional effectiveness». *Journal of Educational Psychology*, 64.
- ANDERSON, C.S. (1982): «The search for school climate: A review of educational research». *Review of Educational Research*, 52, (3) pp. 368-420.
- APARICIO, J, Sanmartín, R y TEJEDOR, J. (1982): *La enseñanza universitaria vista por los alumnos: Un estudio para la evaluación de los profesores en la enseñanza superior*. Cuadernos de Capacitación docente, O.E.I., Madrid.
- BANISTER, J. y otros (1961): «Evaluating college teaching». *Curriculum Reporter Supplement*, 1.
- BARTOLOMÉ, A y MATEO, J. (1983): «Utilización del lenguaje audiovisual en la enseñanza de la estadística». *Enseñanza*, 3, pp. 183-192.

- BEAN, J.P. (1985): «Interaction effects on class level in an explanatory model of college student dropout syndrome». *American Educational Research Journal*, 22, (1), pp. 35-64.
- BORGATTA, E.F. (1970): «Students ratings of faculty». *American Association of University Professors Bulletin*, 56, pp. 6-7.
- BRASKAMP, L.A. (1978): «Colleague evaluation of instruction». *Faculty Development and Evaluation in Higher Education*, 4, pp. 1-9.
- BRASKAMP, L.A. y otros (1979): «Student ratings and instructor self-ratings and their relationship to student achievement». *American Educational Research Journal*, 16, pp. 295-306.
- CENTRA, J.A. (1977): «The how and why of evaluating teaching» en J.A. Centra *Reviewing and evaluating teaching*, Jossey-Bass, San Francisco.
- COSTIN, F. y otros, (1971): «Student ratings of college teaching: Reliability, validity and usefulness», *Educational Research*, 41, pp. 511-535.
- COHEN, P.A. (1982): *Synthesizing research results on teacher evaluation using meta-analytic procedures*. American Psychological Ass., Washington (ERIC Document Reproduction Service n.º ED 223646).
- CHORRO GASCO, J.L. (1977): *Aptitudes mentales y rendimiento universitario*. Tesis Doctoral. Facultad de Filosofía y Ciencias de la Educación. Universidad de Valencia.
- DEMING, W.E. (1972): «Memorandum of teaching». *American Statistician*, 26.
- DOYLE, W. (1977): «Paradigms of research on teacher effectiveness» en L.S. Shulman (ed) *Review of Research in Education*, Itasca, Illinois, Peacock.
- DRUCKER, A.J. y REMMEMRS, H.H. (1951): «Do alumni and students differ in their attitudes toward instructors?». *Journal of Educational Psychology*, 42.
- ESCUDERO, J.M. (1980): «La eficacia docente: estudios correlaciones y experimentales» en Varios, La investigación pedagógica y la formación de profesores. Sociedad Española de Pedagogía. Actas de VII Congreso Nacional. Granada, pp. 207-235.
- FOSTER, D.C. (1984): *Selection of evaluation criteria for the development of a teacher assessment systems*. Washington Uni., Teacher Education research Center (ERIC. n.º ED 242719).
- FREY, P.W. (1978): «A two-dimensional analysis of student ratings of instruction». *Research in Higher Education*, 9, pp. 69-91.
- GILLMORE, G.M. (1975): *Estimates of reliability coefficients for items and subscales of the Illinois Course Evaluation Questionnaire*. Research Report, n.º 341. Urbana, Illinois: Measurement and Research Division Office of Instructional Resources, University of Illinois.
- GIMENO, J. (1981): «La integración de la teoría del aprendizaje en la teoría y práctica de la enseñanza». En A. Pérez Gómez (ed), *Lecturas de aprendizaje y enseñanza*. Zero-Zyx, Madrid, pp. 467-499.
- (1983) «Planificación de la investigación educativa y su impacto en la realidad» en J. Gimeno y A. Pérez Gómez, *La enseñanza: su teoría y su práctica*. Akal, Madrid, pp. 166-187.
- GRUSH, J.E. y COSTIN, F. (1975): «The student as consumer of the teaching process». *American Educational Research Journal*, 12, pp. 55-66.

- GUTHRIE, E.R. (1954): *The evaluation of teaching: A progress report*. Seattle, University of Washington.
- HAYES, J.R. (1971): «Research, teaching and faculty fate». *Science*, 172, pp. 227-230.
- HYNES, C y otros (1985): «Teacher clarity and its relationship to student achievement and satisfaction». *American Educational Research Journal*, Vol. 22, n.º 1, pp. 87-110.
- HOGAN, T.P. (1973): «Similarity of students ratings across instructors, courses and time». *Research in Higher Education*, 1, pp. 149-154.
- KING, J.A. (1981): «Beyond classroom walls: Indirect measures of teacher competence» en J. Millman (ed) *Handbook of teacher evaluation*, Sage Publications, Berwely Hill, California.
- LANDSHERE, G. de (1982): *La investigación experimental en educación*. Unesco, París.
- LINSKY, A.S. y STRAUS, M.A. (1975): «Students evaluations, research productivity and eminence of college faculty». *Journal of Higher Education*, 46, pp. 89-102.
- MARÍN SÁNCHEZ, M. (1983): *Factores y causas del fracaso académico en la Universidad*. ICE, Sevilla.
- MASLOW, A.H. y ZIMMERMAN, W. (1956): «College teaching ability, scholarly activity and personality». *Journal of Educational Psychology*, 47, pp. 185-189.
- MARSH, H.W. (1977): «The validity of students» evaluations: Classroom evaluations of instructors independently nominated at best and worst teachers by graduate seniors». *American Educational Research Journal*, 14, pp. 441-447.
- MARSH, H.W. y OVERALL, J.V. (1979): «Long-term stability of students» evaluations: A note on Feldman's consistency and variability among college students in rating their teachers and courses». *Research in Higher Education*, 10, pp. 139-147.
- MARSCH, H.W. y otros (1979): «Validity of students evaluations of instructional effectiveness: A comparison of faculty self-evaluations by their students». *Journal of Educational Psychology*, 71, pp. 149-160.
- McDANIEL, E.D. y FELDHUSEN, J.F. (1970): «Relationships between faculty ratings and indexes of service and scholarship». *Proceedings of the 78th Annual Convention of the American Psychological Association*, 5, pp. 619-620.
- McGRATH, E.J. (1962): «Characteristics of outstanding college teachers». *Journal of Higher Education*, 33, pp. 148.
- McGREAL, T.L. (1982): *Developing a teacher evaluation system: commonalities of these systems that function most effectively*. Washington, (ERIC, n.º ED 226418).
- McKEACHIE, W. J. y otros. (1978): «A small study assessing teacher effectiveness: Does learning last?». *Contemporary Educational Psychology*, 3, pp. 352-357.
- MCNEIL, J.D. y Popham, W.J. (1973): «The assessment of teacher competence» en R.M. Travers (ed). *Second Handbook of Research on Teaching*. Rand McNally, Chicago.
- MILLMAN, J. (ed) (1981): *Handbook of teacher evaluation*, Sage Publications, Berwely Hill, California.

- MOOMAW, M.T. (1977): «Practices and problems in evaluating instruction» en J.A. Centra (ed). *Renewing and evaluating teaching*, Jossey-Bass, San Francisco.
- NIETO y GIL, J.M. (1982): *Valoración de la eficacia docente. Problemas y Técnicas*. Escuela Española, Madrid.
- ORY, J.C. y otros. (1980): Congruency of student evaluative information collected by three methods». *Journal of Educational Psychology*, 72, pp. 181-185.
- PÉREZ GÓMEZ, A. (1983): «Paradigmas contemporáneos de investigación didáctica» en J. Gimeno y A. Pérez Gómez (ed) *La enseñanza: su teoría y su práctica*. Aka, Madrid, pp. 95-138.
- PERRY, R.P. y otros (1979): «Educational seduction: The effect of instructors expressiveness and lecture content on student ratings and achievement». *Journal of Educational Psychology*, 71, pp. 107-116.
- PETERSON, K. (1984): «Methodological problems in teacher evaluation». *Journal of Research and Development in Education*, 17 (4), pp. 62-70.
- PUGACH, M.C. y Raths, J.D. (1983): «Testing teachers: Analyse and recommendations». *Journal of Teacher Education*, 34 (1), pp. 37-43.
- RAYDER, N.F. (1968): «College students ratings of instrutors». *Journal of Experimental Education*, 37, pp. 76-81.
- RIVAS, F. MONERA, M.L. (1984): «Rendimiento en la Universidad en función del seguimiento o no del consejo vocacional en COU». *Enseñanza*, 2, p. 183-204.
- ROSENSHINE, B. y FURST, N. (1971): «Research in teacher performance criteria» en B.O. Smith (ed), *Research in Teacher Education: A symposium*. Englewood Cliffs, N.J.: Prentice Hall.
- SANDEFUR, J.T. (1983): *Competency testing of teachers*. Washington, (ERIC, n.º ED 238858).
- SCRIVEN, M. (1967): «The methodology of evaluation» en R. Tyler y otros (eds), *Perspectives of curriculum evaluation*, Rand McNally, Chicago. Monograph Series of Cruriculum Evaluation, 1.
- STAKE, R.E. (1967): The countenance of educational evaluation» *Teachers college Record*, 68, pp. 523-540.
- STUFFLEBEAM, D.L. y otros (1971): *Educational evaluation and decision making*. Peacock for Phi Delta Kappa National Study Committe on Evaluation, Itasca, Illinois.
- THEODORY, G.C. y DAY, R.C. (1985): «The association of professors'style, trait anxiety and experience with student's grades». *American Educational Research Journal*, 22 (1) pp. 123-134.
- TOURON, J. (1964): *Factores del rendimiento académico en la Universidad*. Eunsa, Pamplona.
- TUCKMAN, B.Q. y OLIVER, W.F. (1968): «Effectiveness of feedback to teachers as function of source». *Journal of Educational Psychology*, 59, pp. 297-301.
- TYLER, R.W. (1942): «General statement on evaluation». *Journal of Educational Research*, pp. 492-501.

- VILLAR ANGULO, L.M. (1983): «Evaluación de la enseñanza universitaria por los estudiantes». *Enseñanza*, 1, pp. 263-298.
- (1985) *Evaluación del ambiente de aprendizaje en el aula universitaria*. Documento fotocopiado.
- VOECKS, V.W. (1964): «Publications an teaching effectiveness». *Journal of Higher Education*, 33, pp. 212.
- WARE, J.E. y WILLIAMS, R.G. (1977): «Discriminate analysis of student ratings as a means of identifying lecturers who differ in enthusiasm of information giving». *Educational and Psychological Measurement*, 37, 627-639.
- WEIS, C.H. (1972): *Evaluation research: Methods for assesing program effectiveness*. Prentice Hall, Englewood Cliffs, N.J..
- WERGIN, J.F. (1977): «Evaluating faculty development programs» en J.A. Centra (ed) *Renewing and evaluating teaching*, Jossey-Bass, San Francisco.
- WITTROCK, M.C. y LUMSDAINE, A.A. (1977): «Instructional Psychology». *Annual Review of Psychology*, 28, pp. 417-459.
- WORTTMAN, P. M. (1975): «Evaluation research: A psychological perspective». *American Psychologist*, 30, pp. 562-575.

Colección  
**PEDAGOGIA**

Una parte de esta colección la configuran obras de algunos profesores del Departamento de Pedagogía Experimental, Terapéutica y Orientación de la Universidad de Barcelona con probada experiencia en Orientación Educativa y Vocacional.

**BASES TEORICO PRACTICAS PARA LA REALIZACION DEL DIAGNOSTICO PEDAGOGICO:** Ofrece un material para trabajar ordenadamente con alumnos de psicología y pedagogía el proceso de realización del Diagnóstico pedagógico. Contiene esquemas desarrollados de temas y anexos con instrumentos de observación, casos, informes,... (P.V.P.: 1.200 ptas.).

**TEORIAS Y PROCESO DE LA ORIENTACION EDUCATIVA:** Presenta el conjunto de teorías actualizadas de la orientación y el consejo, teorías del concepto de sí mismo, cómo llevar a cabo las entrevistas y cuáles son los principios organizativos de los programas integrales de orientación, además de un panorama de la historia de la orientación desde una perspectiva internacional y de Cataluña. Presenta amplísimas bibliografías. (P.V.P.: 1.200 ptas.).

**MODELOS DE INTERVENCION EN ORIENTACION EDUCATIVA Y VOCACIONAL:** Es un libro fundamentalmente práctico para uso del tutor u orientador que desee organizar algún modelo de orientación a partir de los ciclos superior y medio de la EGB y primeros cursos del BUP. Ofrece amplia bibliografía en castellano para la obtención de recursos en los gabinetes de orientación, y listados de tests idóneos para la orientación profesional. (P.V.P.: 500 ptas.).

**SERVICIOS Y ACTIVIDADES DE INFORMACION PROFESIONAL:** Presenta una visión general de los servicios, las técnicas y actividades de información profesional que se pueden realizar en los centros educativos, teniendo en cuenta las investigaciones y experiencias realizadas últimamente. (P.V.P.: 950 ptas.).

---

Los encontrará en los distribuidores de P.P.U.

**PROMOCIONES Y PUBLICACIONES UNIVERSITARIAS**  
C/. Nicaragua 100, 7.<sup>o</sup>, 1.<sup>a</sup> – 08029 BARCELONA  
Teléf.: (93) 239 91 37

# Comunicaciones

Revista Investigación Educativa - Vol. 3 - N.º 6 - 1985 (P.339-345)

## Rendimiento académico en el área de matemáticas al término del ciclo medio de Enseñanza General Básica

por  
*Justo Arnal*

Es opinión generalizada entre los profesores del área de Matemáticas de los distintos niveles de enseñanza, que a los alumnos suele «costarles» un esfuerzo especial conseguir los objetivos establecidos en los programas oficiales; Es decir, no llegan al nivel de suficiencia (Dirección General de Educación Básica, 1979; Arnal, J. 1984; Vidal, M.C. 1984; Servicio de evaluación e innovación Educativa del Ministerio de Educación y Ciencia, 1984; etc.) (1), dándose un porcentaje alto de fracaso escolar (2) al término de los cursos o de los ciclos escolares de enseñanza general básica.

La reforma de los programas en la enseñanza general básica, introducida en 1981 de forma experimental, y en 1982 ya en forma definitiva, ofrecía la posibilidad de diseñar y realizar un proyecto de investigación evaluativa sobre el área de Matemáticas, a través de la que pudiera obtenerse datos de forma empírica y averiguar si los alumnos que terminaban el ciclo medio en junio de 1984, conseguían el dominio de los objetivos previstos en el área de matemáticas, concretados en los programas renovados.

### FORMULACIÓN DEL PROBLEMA.

El ciclo medio de enseñanza general básica es obligatorio según establece la legislación escolar. El alumno durante tres cursos ha de conseguir los niveles básicos de referencia concretados en los programas renovados para cada área de conocimientos. Al evaluar el nivel de conocimientos adquiridos mediante un test de instrucción válido, externo a la propia clase y redactado de acuerdo con los mencionados niveles, encontramos alumnos con nivel de insuficiente.

## HIPÓTESIS DE TRABAJO

Primera: Los niveles básicos de referencia establecidos para los alumnos del ciclo medio de enseñanza general básica, sólo son conseguidos por la mitad o menos de ellos, si se toma como criterio que el nivel de suficiencia sea el setenta y cinco por cien de la puntuación de un test de instrucción válido y fiable.

Segunda: No existen diferencias de medias significativas , en cuanto a instrucción matemática se refiere, entre chicas que terminan el ciclo medio de enseñanza general básica.

Tercera: Los profesores de E.G.B. tienden a calificar con notas que se encuentran en el intervalo 5-7.

## CARACTERÍSTICAS DE LA MUESTRA

La muestra procedía de veintitrés centros escolares, localizados en catorce provincias españolas situadas en puntos geográficos distintos dentro del Estado: Sevilla, Barcelona, Asturias, Alicante, La Coruña, Murcia, Pamplona, Zaragoza etc. Los centros escolares son privados no confesionales, de nivel socioeconómico medio-alto, situados en núcleos de población urbana.

El procedimiento de selección fue circunstancial o errático. El número de sujetos fue de mil quinientos treinta alumnos de 5.<sup>º</sup> de enseñanza generla básica, de los que 619 son chicas y 911 chicos, muestra equivalente a más de 200.000 sujetos (Krejce y Morgan, 1970).

## INSTRUMENTO DE MEDIDA

Se eleboró el test de instrucción denominado «Mat-50» conseguido a partir del test «Mat-60» (1983) y de «Mat-46» (1984), poseyendo este último una validez de 0,52 y una fiabilidad de 0,81 según el procedimiento de Pearson; 0,89 según Rulon y KR-21 igual a 0,92. Todos estos datos se obtuvieron en una muestra de 2.215 alumnos.

## RECOGIDA DE DATOS

El test de instrucción «Mat-50» se aplicó a finales de mayo y primeros de junio de 1984, siendo los propios profesores quienes realizaron la aplicación en su clase siguiendo las instrucciones escritas que se habían redactado.

**Resultados a partir de las puntuaciones directas totales (n = 1530).**

I	Frecuencia
0- 5	0
5-10	4
10-15	15
15-20	61
20-25	136
25-30	221
30-35	307
35-40	278
40-45	300
45-50	208
	1530

$$\bar{X} = 34,94 ; \quad M_e = 35,38 ; \quad M_o = 37,47 ; \\ S = 8,7 ; \quad A_{s1} = -0,29 ; \quad A_{s2} = -0,15 ;$$

La prueba de normalidad ji cuadrado indicaba el valor de 85,64 y Kolmogorov-Smirnov (D. máxima=0,041). La distribución en estudio se aleja mínimamente de la distribución en estudio se aleja mínimamente de la distribución normal. La fiabilidad mediante KR-20=0,88; y la validez=0,73.

**Distribución resultante a partir de las puntuaciones obtenidas por los alumnos de la muestra.**

I	F
0- 5	0
5-10	2
10-15	5
15-20	38
20-25	76
25-30	111
30-35	182
35-40	151
40-45	187
45-50	159
	911

$$\bar{X} = 35,76 ; \quad M_e = 36,37 ; \quad M_o = 37,13 ; \\ S = 8,6 ; \quad A_{s1} = -0,21 ; \quad A_{s2} = -0,21 ;$$

Según los valores de ji cuadrado (98,97) y Kolmogorov-Smirnov (0,048) la distribución en estudio se aleja de la tenida como normal. La fiabilidad de la prueba con esta muestra fue de 0,88 mediante KR-20, y su validez igual a 0,71.

**Distribución resultante a partir de las puntuaciones obtenidas por las alumnas de la muestra.**

I	Frecuencia
0- 5	0
5-10	2
10-15	10
15-20	23
20-25	60
25-30	110
30-35	125
35-40	127
40-45	113
45-50	49
	619

$$\bar{X} = 33,32 ; \quad Me = 34,18 ; \quad Mo = 32,68 ; \\ S = 8,45 ; \quad A_{s1} = -0,06 ; \quad A_{s2} = -0,34 ;$$

Los valores de las pruebas de normalidad ji cuadrado y Kolmorov-Smirnov son 22,01 y 0,023 respectivamente. La distribución se halla dentro de la normalidad. La fiabilidad resultó ser 0,87 por KR-20 y la validez=0,74.

**Diferencia de medidas**

Del estudio ex-post-facto realizado se deduce una diferencia real entre chicos y chicas:  $35,76 - 33,22 = 2,54$  ¿Esta diferencia es significativa a favor de los alumnos?. La razón crítica era igual a 0,674. Al 95% n.c. la diferencia no es significativa. Esto no quiere decir que no lo sea, sino que no lo es a ese nivel.

**Distribución resultante a partir de las calificaciones dadas por los profesores.**

I	f	%
3- 5	220	14,38
5- 7	714	46,66
7- 9	436	28,49
9-10	160	10,45
	1530	

$$\bar{X} = 6,65 ; \quad Me = 6,53 ; \quad Mo = 6,33 ; \\ S = 1,84 ; \quad A_s = 0,17 ;$$

Los valores de ji cuadrado y Kolmogorov eran 77,83 y 0,0352; la distribución en estudio se aparta de la normal al 95 % n.c.

Análisis de ítems			Resultados en tantos por ciento por temas		
ítem	I.D.	I.V.	Tema	ítem	%
1	0,86	0,22	Conjuntos	1 al 12	69
2	0,71	0,18	Números romanos	13	77
3	0,53	0,17	Medidas de tiempo	14, 15, 16	68
4	0,73	0,34	Potencias	17 y 18	89
5	0,67	0,12	Suma	19	97
6	0,61	0,18	Resta	20	82
7	0,64	0,22	Aplicación suma y resta	25 y 30	75
8	0,39	0,10	División	21	85
9	0,33	0,00	Multiplicación	23 y 24	86
10	0,79	0,12	Aplicación multiplicación y división	45	44
11	0,79	0,12	Fracciones	22, 31, 32, 33 34 y 35	65
12	0,85	0,15	Proporciones	26 y 27	61
13	0,72	0,23	Medidas	28, 36 al 42	67
14	0,54	0,27	Geometría	46 al 50	53
15	0,64	0,23	Media total		73
16	0,70	0,15			
17	0,88	0,23			
18	0,77	0,22			
19	0,90	0,20			
20	0,78	0,18			
21	0,78	0,22			
22	0,74	0,22			
23	0,83	0,27			
24	0,79	0,22			
25	0,67	0,15			
26	0,63	0,27			
27	0,53	0,17			
28	0,67	0,27			
29	0,85	0,32			
30	0,75	0,29			
31	0,52	0,23			
32	0,52	0,24			
33	0,63	0,17			
34	0,57	0,22			
35	0,66	0,24			
36	0,64	0,17			
37	0,59	0,22			
38	0,66	0,12			
39	0,46	0,26			
40	0,56	0,22			
41	0,76	0,24			
42	0,64	0,18			
43	0,47	0,29			
44	0,67	0,17			
45	0,40	0,22			
46	0,19	0,04			
47	0,52	0,17			
48	0,74	0,24			
49	0,53	0,14			
50	0,74	0,20			

## CONCLUSIONES.

Primera hipótesis. No alcanzan el 75% de la puntuación del test 744 alumnos (48,63%), más 83 que se hallan incluidos en el intervalo 35-40; por tanto, 827 alumnos (54,05%) no consiguen el criterio establecido. Se cumple la hipótesis formulada.

Segunda hipótesis. No es significativa la diferencia de medias entre chicos y chicas al 95% n.c. Se cumple la hipótesis formulada.

Tercera hipótesis. Setecientos catorce alumnos (46,66%) se encuentran en el intervalo 5-7. Se cumple la hipótesis.

Conviene observar en la distribución obtenida a partir de las calificaciones dadas por los profesores que el 85,6% de los alumnos alcanzan puntuación positiva, dato que se ajusta al que teóricamente debe alcanzarse (85%). Sin embargo, 220 alumnos (14,38%) quedan con nivel insuficiente; pensamos que son demasiados.

En la distribución resultante de los 1530 alumnos a partir de las puntuaciones obtenidas en el test de instrucción encontramos 528 sujetos próximos a la media muestral y éstos bien pudieran formar un subconjunto sobre los que los profesores han «volcado su generosidad» (evaluación insuficiente pero satisfactoria) al dar la calificación definitiva final de ciclo.

En aprendizajes instructivos básicos, como son las matemáticas en EGB, consideramos que el nivel de suficiencia debería situarse en el dominio del 70-80% de los niveles de referencia (Programas Renovados). Caso de que un alumno no alcance una puntuación próxima a éste criterio, tal vez conviniera realizar el correspondiente diagnóstico y reorientar su aprendizaje de forma personalizada a través de un «material de apoyo» que le permitiese seguir con éxito el próximo curso.

La geometría aparece como una de las partes marginadas y convertida en cienfienta del área de matemáticas junto con el bloque temático de conjuntos.

## NOTAS Y REFERENCIAS BIBLIOGRÁFICAS

- (1) DIRECCIÓN GENERAL DE EDUCACIÓN BÁSICA (1979): *Resultados y rendimiento del sistema escolar*. Dirección General Básica. Servicio de Inspección Técnica (EGB). Madrid.
- ARNAL, J. (1984): «Niveles básicos de referencia del área de matemáticas del ciclo medio en Barcelona ciudad y provincia», comunicación presentada a las *II Jornadas de estudio sobre investigación en la escuela*, Sevilla 7-9 de diciembre, págs. 148-151.
- VIDAL, M.C. (1984): *Criteris d'Avaluació*. Ceac. Barcelona.
- MINISTERIO DE EDUCACIÓN Y CIENCIA (1984): *Evaluación de enseñanza mínima en el ciclo inicial de Enseñanza General Básica*. Servicio de Evaluación e Innovación Educativa del MEC. Madrid.

- 
- (2) Entendemos aquí por fracaso escolar no conseguir los objetivos fijados en los programas renovados en su gran mayoría.

DOSSEY, J y otro (1984): *Summary of the 1982 Mathematics Results of the Illinois Inventory of Educational Progress*. Illinois State Board of Education, Springfield. Many.

MARELIM, H. (1983): *Los tests en la educación*. Eunsa. Pamplona.

RAMOS, A. (1982): «Programas renovados de matemáticas. Ciclo medio». *Bordón* núm. 242-243, marzo-junio, págs. 265-281.

REAL DECRETO 710/1982, de 12 de febrero, por el que se fijan las enseñanzas mínimas para el ciclo medio de E.G.B. (B.O.E. de 15 de abril).

ORDEN MINISTERIAL de 6 de mayo de 1982 (B.O.E. de 14 de mayo).

RESOLUCIÓN DE LA DIRECCIÓN GENERAL DE EDUCACIÓN de 29 de septiembre de 1982, regula la evaluación en el ciclo medio. *Vida escolar*, núm. 220-221, septiembre/diciembre, 1982.

GENERALITAT DE CATALUNYA. Departament D'Ensenyament (1982): «Orientaciones i programes bàsics de cicle mitjà d'EGB. Full D'Informació a les escoles, especial, n.º 19, setembre.

GENERALITAT DE CATALUNYA. Departament D'Ensenyament (1982): Cicle Mitjà. L'avaluació al Cicle Mitjà. *Full D'Informació a les escoles, Extraordinari*, núm. 22, novembre-diciembre.

GENERALITAT DE CATALUNYA. Departament D'Ensayament (1983): Orientacions i programes. Cicle Mitjà D'Educació General Básica.

GENERALITAT DE CATALUNYA. Departament D'Ensenyament (1985): «Monográfic Cicle Mitjà.» *Butlletí dels mestres*, n.º 197, agosto-septiembre, Barcelona.

# Análisis comparativo de formas verbales e icónicas en la medición del rendimiento

por  
*Antonio R. Bartolomé*

## INTRODUCCIÓN

El uso de elementos icónicos en ítems en diferentes tipos de pruebas y tests no es novedoso. Sin embargo, en las pruebas utilizadas habitualmente en los diferentes niveles educativos no es habitual el uso de figuras. En un análisis sobre 4 pruebas de rendimiento tipo test utilizadas en la asignatura «Estadística aplicada a las Ciencias de la Educación» menos del 5% de los ítems incluían elementos icónicos.

¿Cómo es que el uso masivo de mensajes icónicos en el proceso de aprendizaje no se corresponde con un uso similar en el momento de la evaluación? Apuntamos dos explicaciones:

- a) a nivel universitario la imagen tiene un papel subordinado a la comunicación verbal;
- b) un mayor uso de imágenes en las pruebas de rendimiento no alteraría sustancialmente los resultados académicos; esto junto con el mayor costo de realización haría poco rentable su uso.

En este trabajo se estudia esta segunda hipótesis.

## ALGUNAS REFERENCIAS TEÓRICAS

La presentación de un ítem en forma verbal o icónica no es indiferente. La relación entre imagen y pensamiento ha sido abordada desde una perspectiva piagetiana (Sinclair, H. 1978), Psicométrica (Guilford, J.P. 1977) o desde una concepción de la inteligencia como «procesadora de la información» (Wagner, R.K. y Sternberg, R.J. 1984). Recientes trabajos en Psicología de Gringer y

Bandler (Patton M.Q. 1981) sugieren que personas diferentes han desarrollado diferentes sistemas (verbales, tátiles, sonoros y visuales) de representación e interpretación del mundo. M.Q. Patton concluye que no hay una modalidad de comunicación y experiencia única y efectiva para todos.

La imagen *no* siempre ayuda a resolver correctamente un ítem. Experiencias de Wohlwill y el trabajo de Winer y Kronberg (Denis, M. 1984) recogen algunos ejemplos en esta línea.

En nuestro caso, la introducción de elementos icónicos está referida a la codificación efectiva de información fundamental del ítem mediante códigos visuales. No se incluye el uso del comic como elemento motivador, siguiendo en ese sentido a Macdonald-Ross, M. (1978).

## PLANTEAMIENTO DE LA INVESTIGACIÓN

El instrumento utilizado fue un test elaborado desde la perspectiva de una evaluación basada en criterios, concretamente siguiendo los planteamientos de Popham W.J. (1980, 1983). A partir de unas especificaciones se establecieron para cada uno unos contenidos equivalentes: sobre este material se elaboraron ítems apareados, uno en forma verbal y otro en forma icónica. (*figura 1*).

Los ítems fueron asignados aleatoriamente-apareados a dos modelos (A y B) de la prueba. Cada modelo recogía ítems correspondientes a las distintas especificaciones, pero para una misma especificación, los ítems pertenecían a diferentes contenidos para evitar fenómenos de resonancia entre las formas icónica y verbal.

Los dos modelos se distribuyeron aleatoriamente según un esquema sistemático sobre la distribución espacial. La muestra era un grupo académico de 46 estudiantes de 2.<sup>º</sup> curso de Pedagogía. También se aplicó a este grupo una prueba clásica de problemas.

La muestra no tiene carácter de representatividad. Se trata de un estudio puntual destinado a ser replicado y mejorado en diferentes situaciones. Las técnicas inferenciales utilizadas en el análisis son tales que permiten utilizarlo como elemento metodológico en clase con los mismos alumnos.

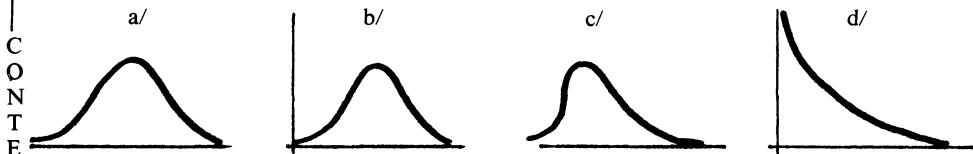
## ELABORACIÓN DE LA PRUEBA

El test pretendía medir un dominio preciso: la aplicación por parte de los estudiantes de técnicas inferenciales de estimación y contraste en base a las distribuciones Normal, t-Student y ji-cuadrado. La prueba se construyó según muestra el siguiente esquema:

cuatro pruebas tipo test  
utilizadas otros años  
(análisis de contenido)

fig. 1

5. ¿Cuál de entre los siguientes modelos de distribuciones de probabilidad sigue la distribución muestral de una proporción?



I D O 5. ¿Cuál de entre los siguientes modelos de distribuciones de probabilidad sigue la distribución muestral de una porción?

- I a/ Ley Normal
- b/ Ley Normal asimétrica
- c/ distribución t con 1 grado de libertad
- d/ distribución chi-cuadrado

II C O N T E N I D O 13. ¿Cuál de entre los siguientes modelos de distribuciones de probabilidad sigue la distribución muestral de una media?



II C O N T E N I D O 13. ¿Cuál de entre los siguientes modelos de distribución de probabilidad sigue la distribución muestral de una media?

- a/ Ley Normal asimétrica
- b/ t de Student con x grados de libertad
- c/ chi-cuadrado
- d/ F de Snedecor

1. Con referencia a las *distribuciones muestrales* correspondientes a
    - . proporción en que un carácter se encuentra presente
    - . distribución de sujetos entre varias categorías
    - . media de una variable cuantitativa
- identificar:*
- . modelo de probabilidad que sigue

listado de 16 destrezas

(reagrupación; comparación con objetivos de la asignatura)  
listado de 13 destrezas básicas

(encuesta valorativa entre 4 profesores de la asignatura)

#### ESPECIFICACIONES

(redacción a partir de especificaciones y temario)  
(evaluación y corrección por jurado de profesores del Departamento)

Contenidos alternativos de dificultad presuntamente equivalente

(elaboración de items por parejas, con formas icónicas y verbales)

#### ÍTEMES

(asignación aleatoria por parejas)

#### MODELOS A y B

La prueba fue fiabilizada por un procedimiento similar al sugerido por Popham W.J. (1983, pg. 206). El criterio de decisión se estableció en un 70% de aciertos («experto»). Se obtuvo un coeficiente estimado entre .551 y .601 (alfa = .05). Al existir dos modelos (A y B) también se compararon ambos. Dado el carácter del test se prefirió una prueba no-paramétrica (U de Mann-Whitney), obteniendo zeta = .76 (grado de signif. aprox. .45).

La validez descriptiva se trabajó en el proceso de elaboración de las especificaciones; este tipo de validez es el más importante para la medición basada en criterios, según Mehrems W.D. (1973) y fue especialmente cuidada. Desde un punto de vista funcional se estudió la correlación entre la calificación («experto», «no-experto») en el test y en una prueba clásica de problemas. Se obtuvo un valor Phi = .323 (significativo al nivel .05). Estudiando las fuentes de error debidas al diferente carácter de ambas pruebas se señalan:

- a) existencia de cálculos en la prueba de problemas
- b) posibilidad de aciertos al azar en el test.

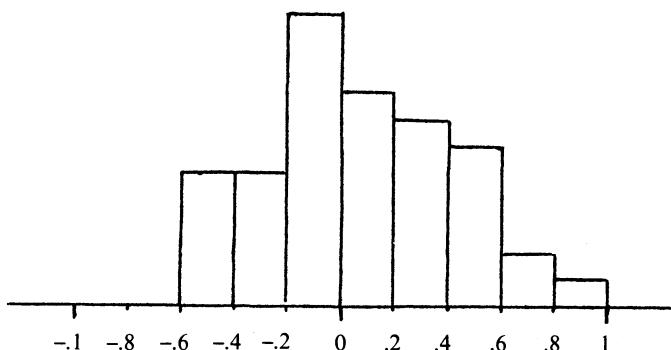
Notemos que en ambos casos se tiende a incrementar el error tipo II, y a reducir el valor numérico de Phi.

No nos detenemos más en el análisis de los ítems, remitiendo a lo que indica Popham W.J (1983, pg. 140) sobre la necesidad de una homogeneidad derivativa y no una funcional. Constatemos de cualquier modo la escasez de mecanismos eficaces y seguros en este terreno, al menos por comparación con los tests basados en normas (Haertel, E. 1985).

## ANÁLISIS DE RESULTADOS

Inicialmente se calcularon los coeficientes Phi entre los aciertos en una u otra forma (íónica-verbal) para cada individuo de la muestra. Sólo en 3 de los 46 se encontraron coeficientes significativos (*figura 2*). A la vista de estos resul-

Figura 2



(coeficientes PHI para cada individuo de la muestra entre la calificación obtenida en ítems icónicos y la obtenida en verbales)

tados podría pensarse que existen diferencias entre el modo como responden los alumnos en uno u otro caso. Sin embargo conviene hacer dos matizaciones.

En primer lugar, siguiendo a Popham, en un test referido a criterios el peso se desplaza del ítem al conjunto de la prueba. Así no interesa tanto constatar si cuando un alumno acierta un ítem en forma icónica, acierta también otro similar presentado en forma verbal, sino si los alumnos que serían considerados expertos en base al conjunto de los ítems verbales, también lo serían en base a los ítems icónicos. Aplicando un coeficiente Phi para el conjunto de la muestra según lo indicado obtenemos .5167 (significativo al nivel .05). En el 77% de los casos un alumno obtenía la misma calificación basándose en los ítems icónicos o en los verbales.

En segundo lugar recordemos que cuando apareamos los ítems correspondientes a una misma destreza, para cada alumno, no correspondían únicamente a diferentes formas (verbal o icónica) sino también a diferentes contenidos (supuestamente equivalentes). Por ello se procedió a analizar los índices de dificultad de los ítems según destrezas, contenidos y formas. A partir de estos datos se contrastaron las proporciones de aciertos entre formas (dentro de un mismo contenido) y entre contenidos (dentro de una misma forma). Se encontraron 4 diferencias significativas en el primer caso, y 9 en el segundo (nivel .05). Las diferencias entre contenidos se acentuaron cuando los ítems los presentaban en forma verbal (7/10 diferencias significativas).

Los ítems en los que se encontraron diferencias significativas por algunos de los conceptos anteriores fueron analizados, obteniéndose algunos resultados interesantes. Alguno se ha incluido en las conclusiones.

El índice de aciertos general fue de .61 para la forma icónica y .63 para la verbal. La distribución de aciertos entre diferentes destrezas, tanto para los contenidos I como los II, no mostraron diferencias significativas entre las formas (grados de significación de 0.50 y 0.80).

## Conclusiones

Para el conjunto de la prueba, las diferencias encontradas no afectan sustancialmente los resultados, si bien invitan a seguir profundizando en la causa de las diferencias.

En ítems concretos se encuentran diferencias muy significativas que sugieren que las especificaciones normalmente deberían incluir indicaciones sobre la conveniencia de utilizar imágenes o no en el ítem. Por otro lado, la misma presentación de los textos o datos actuó como una «imagen» alterando la percepción del ítem.

La falta de homogeneidad entre contenidos supuestamente equivalentes lleva a destacar la eficacia del azar para elaborar pruebas equivalentes (por ejemplo, a partir de un banco de ítems) frente al clásico sistema de elaborar similares. Parece que la forma icónica suaviza las diferencias entre contenidos.

En esta línea podría estudiarse si las formas verbales discriminan más, y si esa discriminación es relevante.

En próximos años, esta prueba se aplicará a alumnos que utilizarán intensivamente el vídeo, lo que permitirá estudiar los cambios que se produzcan en los resultados.

## REFERENCIAS BIBLIOGRÁFICAS

- DENIS, M.: (1984) *Las imágenes mentales*. Siglo xxi, Madrid.
- GUILFORD, J.P.: (1977) *Way beyond the I.Q.* The Creative Education Foundation Inc. Buffalo.
- HAERTEL, E.: (1985) «Construct Validity and Criterion-Referenced Testing», en *Review of Educational Research*, vol. 55, N.<sup>o</sup> 1, pp. 23-46.
- MACDONALD-ROSS, M.: (1978) «Graphics in texts», en *Review of Research in Education* 5, pp. 49-85. F.E. Peacock Publishers, Itasca (ILlinois).
- MEHRENS, W.A.: (1973) *Measurement and Evaluation in Education and Psychology*. Holt, Rinehart and Winston, New York.
- PATTON, M.Q.: (1981) *Creative Evaluation*. Sage Publ., London.
- POPHAM, W.J.: (1980) *Problemas y técnicas de la evaluación Educativa*. Anaya, Madrid.
- POPHAM, W.J.: (1983) *Evaluación basada en criterios*. Edit. Magisterio Español, Madrid.
- SINCLAIR DE ZWART, H.: (1978) *Adquisición del lenguaje y desarrollo de la mente*. Oikos Tau. Barcelona.
- WAGNER, R.K. y STERNBERG, R.J.: (1984) «Alternative Conceptions of Intelligence and their implications for Education» en *Review of Educational Research*, vol. 54, n.<sup>o</sup> 2. Washington.

# **Variables familiares, preescolarización y rendimiento en el primer curso de EGB**

*por  
José Benito*

¿Cómo contribuyen las variables socio-económico-culturales familiares a la determinación de los niveles formativos y cognoscitivos de los niños que inician la EGB? ¿Cómo lo hacen respecto al rendimiento escolar cuando finaliza el Curso primero? Las variables familiares, maduración formativa, nivel cognoscitivo, ¿se comportan de la misma manera en muestras de preescolarizados que en muestras de niños que no acudieron a las aulas del parvulario? La presente Comunicación intenta responder a estas preguntas reanalizando algunos de los datos que nos proporciona la ambiciosa y compleja investigación que está llevando a cabo De Miguel, M. (1979).

Dos muestras, una de sujetos preescolarizados ( $N = 547$ ) pertenecientes a toda la gama de niveles familiares socio-económico, culturales, etc., otra de No-Preescolarizados ( $N = 193$ ) fueron evaluados en diversos aspectos del ámbito cognoscitivo-instrumental (lenguaje, lectura, dominio de cantidades y número, etc.), de la maduración formativa (Hábitos de observación y reflexión, de trabajo y orden responsable, de convivencialidad normada, etc.) al iniciar el primer curso de escolaridad obligatoria. Designaremos estas variables con las siglas «GC» y «RS», respectivamente. Al fin del primer año de EGB les fueron aplicados tests normalizados, de entre los que aludiremos tan sólo a los de Thorndike-Hagen-Lorge «Aptitudes cognoscitivas primaria I», bajo la sigla «Ap C». La nota asignada por el Profesor como expresión general del rendimiento logrado en curso primero será denominada «CGP» (Calificación Global del Profesor).

## **Variables familiares y formación en el inicio de la EGB.**

De Miguel, M. (1979) demuestra detalladamente que las medias generales de los niños preescolarizados son superiores a las de No-Preescolarizados. Sin embargo, encuentra pruebas de que la «Preescolarización no llega a paliar los efectos de «clase».

Por ej., la media en rendimiento de escolares preescolarizados pertenecientes al nivel «Obreros no-cualificados» es inferior en el primer curso a la de niños de «Obreros cualificados» que no habían cursado preescolar.

Es, pues, interesantes estudiar cómo se comporta el sistema de variables en cada una de las muestras, a fin de completar con una perspectiva «causal» los análisis de las diferencias inter-medias. Uno de los enfoques posibles nos lo proporciona el Path-analysis. Por mor de la brevedad y para facilitar la representación gráfica, condensamos en cuatro variables las siete primitivas de este trabajo, efectuadas las correcciones estadísticas oportunas.

VF = Variables familiares, condensación de las originales de M. de Miguel «Acervo cultural familiar», «Clima afectivo familiar» y «Ayuda familiar» conferida al niño en los años que corresponden a las edades del Preescolar.

RS = Reflexividad, Sociabilidad antes aludidas, que es otra variable compleja, asimismo condensada.

GC Calificación global cognoscitiva al iniciarse el primer curso.

CGP = Es la calificación del Profesor al término de Primero.

Es razonable suponer que el orden en que actúan sea el mismo que el orden en que acaban de ser expuestas. Que los datos se ajustan a las condiciones requeridas por el Path-analysis puede ser defendido, pero aquí lo damos por sentado.

¿Afectan todas las variables a aquellas que la hipótesis supone posteriores? ¿Y en qué cuantía? Calculados los coeficientes Path, los resultados pueden representarse según lo hacen las figuras 1 y 2.

Figura 1  
NO PREESCOLARIZADOS

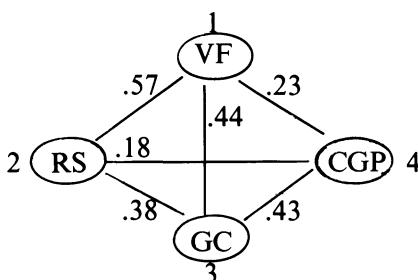
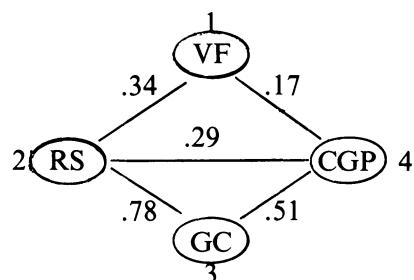


Figura 2  
PREESCOLARIZADOS



Incluso una representación tan simple proporciona informaciones llenas de interés. En primer lugar, los factores antecedentes o causales de tipo familiar (VF) no afectan directamente al nivel cognoscitivo (GC) con el que el niño comienza de EGB, cuando aquél permanece escolarizado de los cuatro a los seis años, mientras que en los No-Preescolarizados el coeficiente Path  $P_{42} = 0,44$ . En segundo lugar,  $P_{21}$  y  $P_{32}$  parecen considerablemente distintos. El primero evalúa la influencia directa de VF sobre los aspectos formativos reunidos bajo

RS. El segundo, cuantifica la intensidad del influjo directo de RS sobre la formación global instructiva. La maduración mental determina el nivel instructivo de los niños preescolarizados en mayor cuantía de lo que lo hace en los No-Preescolarizados. Es en estos mucho más probable un desfase entre sus capacidades socio-mentales y su nivel instructivo. Lo que si por una parte puede no sorprender, ya que el No-Preescolarizado, por el hecho obvio de no haber sido institucionalmente instruido, no ha podido utilizar su capacidad en adquirir las destrezas que proporciona la escolarización, por otra nos está indicando que parte de esas capacidades están siendo desaprovechadas cuando crece en su hogar, sin acudir a la escuela.

Pero GC está ligado a VF en esta muestra no sólo por un nexo directo, sino también a través de la relación entre VF y RS. Es oportuno preguntarse: ¿Qué proporción de la varianza de GC se debe a RS cuando prescindimos de la influencia indirecta de VF? Matemáticamente, ello equivale a hallar la correlación semiparcial entre RS y GC, controlando a VF respecto a la primera.

Haciéndolo así, descubrimos que no llega a 10% la varianza común entre RS y GC en los no preescolarizados ( $R^23(2.1)=0,0959$ ), mientras que se eleva a más del 50% en la otra muestra ( $R^23(2.1)=0,525$ ). Si decidíramos prescindir totalmente de VF, eliminando su influjo tanto de RS como de GC, la proporción de varianza común a RS y GC sería de 0,168 en los No-Preesc. y de 0,574 en los Preescolarizados. De otra manera: si todas las familias (en cada muestra) fuesen igualmente efectivas en el moldeamiento formativo y en la instrucción de destrezas básicas, de suerte que la variabilidad existente entre las mismas se anulara cuanto a Acervo cultural, Clima afectivo y Ayuda prestada al niño, el porcentaje de varianza de GC debida a RS sería 57% en los que terminan Preescolar, cuando no llegaría el 17% en los No-Preesc. (Estas cifras no podrían haberse previsto si tan sólo hubiéramos considerado las correlaciones simples entre RS y GC).

¿Qué acontece con la relación entre VF y RS? En las figuras anteriores, el coeficiente  $P_{21}$  es reflejo de las correlaciones entre los constitutivos de estas variables complejas (los límites de confianza no se solapan aún al nivel 0,01). Descompongamos VF en sus componentes (Ac = Acervo familiar cultural; Af = Clima afectivo; Ay = Ayuda familiar). Por «GF» entenderemos la maduración formativa tal como fue evaluada por el Colegio al dar comienzo la EGB (y en ella se incluyen más aspectos que los únicos R y S hasta ahora tenidos en cuenta. Cfr. la investigación original a la que esta comunicación remite). ¿Cómo contribuyen ac, Af y Ay a la determinación de GF? Sin necesidad de exponer un nuevo gráfico de Análisis de Sendas, baste con decir que Ac y Af fueron consideradas variables primigenias, de las que dependían Ay y GF, en este orden. En la Tabla 1 se exponen los coeficientes Beta, las correlaciones semiparciales –según el orden indicado– y la correlación múltiple de las tres con GF.

TABLA 1

Coeficientes Beta, correlaciones semiparciales y correlaciones múltiples de Ac, Af, Ay con GF en las dos muestras.

		1 Ac	2 Af	3 Ay	$R_{GF.123}$
Prees.	Betas	.188	.135	.093	
	SP	.210	.155	.090	.276
No-Prees.	Betas	.225	.301	.255	
	SP	.400	.354	.221	.578

Un examen atento de la Tabla 1 pone de relieve que, una vez más, los Preescolarizados dependen menos de su formación global de cada uno de los aspectos de VF, y del conjunto de los mismos. Lo que sugiere que los desniveles que fueran debidos a la desigual eficacia educadora de las familias pueden ser compensados por la colegiación de los niños durante los años de preescolaridad.

### VARIABLES FAMILIARES, FORMACIÓN INICIAL Y RENDIMIENTO FINAL EN 1.<sup>º</sup> EGB

Pero ¿cómo intervienen las variables que manejamos respecto al rendimiento final del primer curso? La varianza común a CGP (Calificación global del Profesor) y a las variables RS y GC es mayor en los que hicieron preescolar que en los niños de la otra muestra: en efecto, en los primeros  $R^2 = 0,669$ . En los No-Preesc.  $R^2 = 0,526$ .

(Los intervalos fiduciales de ambos coeficientes comienzan a solaparse a nivel inferior a 0,05). Pero más interesante es preguntarse en qué relación se hallan RS y GC con CGP cuando las variables familiares no contribuyan en nada a la variabilidad del rendimiento final. Esta es en parte debida a aquellas (recuérdese la flecha que liga directamente a VF con CGP en las figuras 1 y 2). Pero si congeláramos el influjo de VF en el rendimiento final, éste se hallaría sólo determinado por la variabilidad en los niveles de maduración y preparación cognoscitiva con que el 1.<sup>º</sup> se inicia, es decir, por lo que fue adquirido por el niño, bien en el ambiente extraescolar en el que maduró y se desarrolló sin acudir al Colegio, bien en el Preescolar. Podríamos así, indirectamente, estimar cuánto del estado inicial del sujeto contribuye a fijar la dispersión que reflejan las calificaciones del rendimiento dadas por los Profesores de primero. Es decir, en qué medida el Preescolar como tal contribuye a los resultados finales.

En términos estadísticos se trata de hallar un coeficiente de correlación múltiple-parcial entre CGP y RS, GC, cuando se deja a VF sin efectos en la variabilidad de la primera. He aquí el resultado:

$$\begin{array}{ll} \text{NO PREESCOLARIZADOS} & R^2 4(23).1 = 0,276 \\ \text{PREESCOLARIZADOS} & R^4 4(23).1 = 0,630 \end{array}$$

En los Preescolarizados, como se ve, la variabilidad existente al final del curso Preescolar prefigura estrechamente la que se obtendrá al final del primer curso, cuando se elimina de ésta el efecto diversificador de las Var. Familiares. Mucho más estrechamente que el estado inicial del niño procedente de su hogar.

Pero de hecho, como sabemos, la correlación entre VF y CGP (ya se ejerza directamente, ya a través de RS y CG) dista de ser despreciable. Si reintroducimos Var. Familiares, sin anular su efecto, como hace un momento, la correlación múltiple al cuadrado de CGP con las tres antecedentes se eleva a 0,554 en los No-Preesc., mientras que en el grupo de Preesc. sólo ascendemos de 0,630 a 0,694.

Descompongamos VF en sus componentes. ¿Cuál es su contribución al rendimiento final? Conservando el orden de entrada que atrás se indicó, la respuesta se resume en la siguiente Tabla:

TABLA 2

Correlaciones semiparciales al cuadrado de las tres variables familiares con CGP

		1 Ac	2 Af	3 Ay	$R^2_{CGP,123}$
Preesc.	SP <sup>2</sup>	.073	.082	.026	.181
No-Preesc.	SP <sup>2</sup>	.160	.140	.095	.395

Todos los caminos llevan a la misma conclusión: los No-Preesc. se hallan condicionados por los factores socio y psico-ambientales familiares, en su rendimiento final, mucho más de lo que lo están los niños que fueron escolarizados en las edades del parvulario. En nada de lo anterior se hace referencia a niveles adquiridos, es decir, a diferencias de logros en el rendimiento entre los dos grupos, puesto que sólo hemos manejado correlaciones y, por tanto, covarianzas y, aceptadas las hipótesis del Path-analysis, nexos causales e intensidad de los mismos. Pero la gran dependencia respecto a VF de los No-Preesc. tanto en el inicio de la EGB cuanto, como acabamos de ver, al término de Primero, nos lleva a la certidumbre de que si estos niños proceden, como sucede en general, de medios familiares y sociales empobrecidos, el Curso primero no compensará las privaciones formativas con las que comienzan su vida escolar. Y aun si el niño ha

sido efectivamente escolarizado, y la educación Preescolar no ha sido planeada cuidadosamente, es muy posible que, en las franjas extremas, las flechas que en la figura 2 apuntan a RS y a CGP estén señalando también a un subrendimiento de recuperación posterior difícil.

#### REFERENCIAS BIBLIOGRÁFICAS

- DE MIGUEL, M. (1979): *Diseño de un programa de Educación Commensatoria en función de los determinantes del primer ciclo de E.G.B.*, I.C.E., Oviedo.

# Elaboración de la POLC (Prueba Objetiva de Lengua Catalana)

por  
*Rafael Bisquerra*

## INTRODUCCIÓN

La reciente incorporación de la enseñanza de la lengua catalana en los niveles obligatorios en Cataluña presenta, entre otras, unas características peculiares que la diferencian de otras asignaturas. Destaquemos la ausencia de unos programas detallados en segunda etapa de EGB i unos grandes desniveles entre distintos centros en BUP.

En todas las materias es un paso importante la evaluación. Dadas las circunstancias actuales, la necesidad de una evaluación objetiva de la lengua catalana es primordial. Por eso, con este trabajo se pretende elaborar y experimentar una Prueba Objetiva de Lengua Catalana (POLC), con tres formas paralelas, con una interpretación referida al criterio y otra referida a la norma.

## FASES EN LA ELABORACIÓN DE LA POLC

En la elaboración de la POLC se han seguido los pasos habituales, expuestos en los principales manuales sobre medida y evaluación, y que se resumen seguidamente.

### 1. Formulación de una propuesta de objetivos de Lengua Catalana

Siguiendo las sugerencias de Rodríguez Diéguez, J.L. (1980) se han seguido los pasos que vienen esquematizados en el Diagrama 1 para la formulación de los objetivos que originarán la prueba.

En la selección de contenidos se tuvieron en consideración: a) las propuestas de programación presentadas por la Comissió Mixta Ministeri d'Educació i Ciència-Generalitat de Catalunya (1979a, 1979b), y Maña, T y Ribas, T (1982).

- b) Los libros de texto de las principales editoriales (Bruño, Casals, Edicions 62, EMESA, Grup Promotor/Santillana, Teide, Vicens Vives), a partir de los cuales se realizó un análisis de contenido.
- c) Las gramáticas de la lengua catalana más usuales (Badia Margarit, Fabra, Jané, Martí i Castell, Marvà, Moll, Rafel i Fontanals y Ruaix i Vinyet).

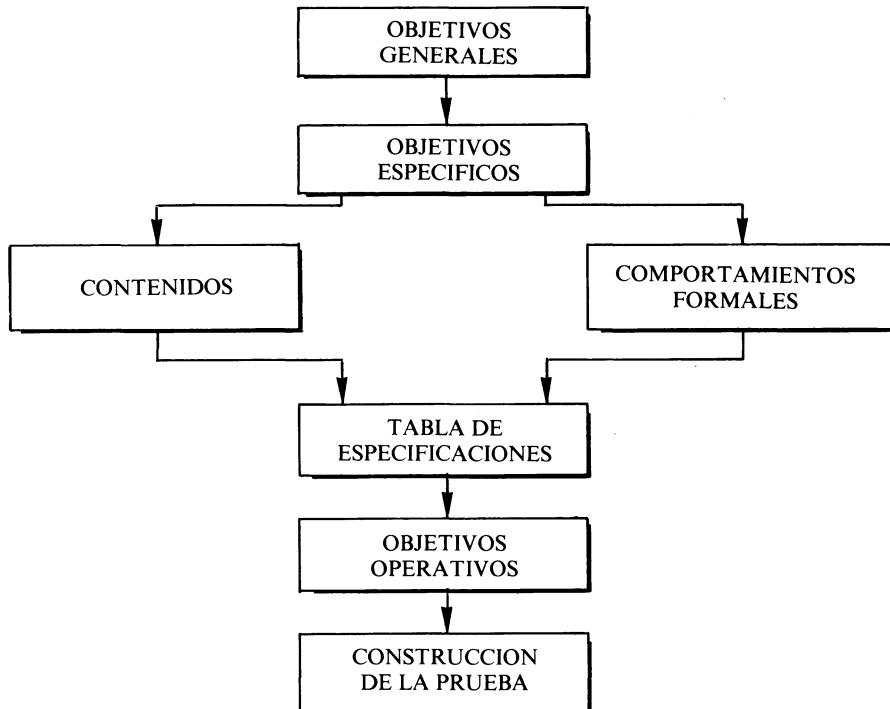


Diagrama 1.-Formulación de los objetivos

Como resultado de este trabajo se presentaron cuatro grandes bloques de objetivos: 1) sociolingüística i dialectología; 2) fonética y ortografía; 3) morfología y sintaxis; 4) léxico y semántica. Cada uno de estos bloques presenta diversos apartados, cada uno de los cuales tiene formulados varios objetivos operativos hasta un total de más de cuatrocientos. Para la selección de los comportamientos formales se ha seguido la taxonomía de Bloom (1975).

Es conveniente tener presente que los objetivos propuestos se refieren a aquellos aspectos que un habitante de Cataluña debería dominar para comprender, hablar, leer y escribir correctamente en catalán, independientemente del nivel educativo que se considere. Por eso se pretende que la POLC sea aplicable a partir de finales de 8.º de EGB, y a lo largo de todo el BUP-COU y FP, así como entre adultos, estableciendo para cada caso las interpretaciones «ad hoc».

## 2. Construcción de los ítems

En la construcción de los ítems se han tenido presentes los siguientes criterios:

- a) Cada ítem responde a un objetivo operativo, a fin de asegurar una validez de contenido de la prueba.
- b) Se han elaborado grupos de tres ítems de características idénticas, a fin de originar tres pruebas paralelas.
- c) Se elabora un 20% de ítems superior a lo necesario, con objeto de eliminar los que después de una primera aplicación no demuestren un suficiente grado de discriminación.
- d) La rapidez y facilidad, tanto en la aplicación como en la corrección, son factores primordiales en la elección del tipo de ítems. También se ha pensado en la posibilidad de corrección mecanizada a través de ordenador. Por eso se ha elegido el ítem de elección múltiple con cuatro alternativas como el más idóneo para satisfacer todos estos requisitos.
- e) Se han seguido las recomendaciones presentadas en los principales manuales.

En total se elaboraron 360 ítems.

## 3. Aplicación piloto de la POLC

Con objeto de someter a experimentación los primeros 360 ítems se seleccionó una muestra de 484 estudiantes de BUP. Característica común a todos ellos era haber cursado enseñanzas de catalán desde los primeros niveles de EGB.

A partir de los datos obtenidos se realizó un análisis de ítems y se eliminaron los 60 con un índice de discriminación inferior, procurando mantener de todas formas grupos de tres ítems de características similares que fueron distribuidos entre las tres formas paralelas de que consta la POLC.

## 4. Aplicación normativa de la POLC

Las tres formas paralelas resultantes fueron administradas a una muestra de 1415 alumnos de toda Cataluña. Los niveles educativos abarcaban desde 8.<sup>º</sup> de EGB hasta COU, incluidos los tres cursos de BUP. Si bien por múltiples causas la muestra no pudo ser aleatoria (hay que tener presente que un requisito era tener cursadas asignaturas de catalán desde principios de la EGB), la muestra puede considerarse que ofrece garantías de representatividad para una primera aproximación al problema.

Esta aplicación experimental tuvo lugar entre el 15 de mayo y el 15 de junio de 1984.

## ANÁLISIS DE RESULTADOS

a) La fiabilidad de la POLC, calculada por el método de Kuder-Richardson, oscila entre el .81 y .91, según formas y niveles.

b) Además de la validez de contenido, dada por el proceso de construcción de la prueba, se ha calculado la validez concurrente tomando como criterio las notas finales de los sujetos en la asignatura de catalán. Los resultados oscilan entre .54 y .70 según formas y niveles.

Se ha calculado también la correlación entre la POLC y otros instrumentos de medida del catalán: dictados, composición escrita y comprensión lectora. Los coeficientes de correlación hallados oscilan entre .38 y .70, variando según formas, niveles y criterio.

c) El error típico de medida, calculado por  $\sigma_m = \sqrt{1 - r^2}$ , oscila entre 6,01 y 6,62, según formas y niveles.

d) El análisis de la dificultad de los ítems demuestra la presencia de ítems de dificultad media en su mayoría, con una cierta cantidad de fáciles y también de difíciles. El índice de dificultad total oscila entre el .41 y el .68, según formas y niveles.

e) El índice de discriminación, calculado por medio de la correlación biserial puntual entre las notas y el acierto o no del ítem, es superior en todos los casos a .32.

f) Se ofrecen unos baremos provisionales para los niveles de: 8-EGB, 1-BUP, 2-BUP, 3-BUP, COU, tanto para puntuaciones directas como para puntuaciones corregidas de la influencia del azar mediante la fórmula  $A - \frac{E}{3}$  donde A = aciertos y E = errores; 3 es el resultado de  $n-1$ , siendo n el número de alternativas de cada ítem.

g) Cuando se elabora un test con varias formas paralelas, se pretende de que éstas sean equivalentes. Sin embargo se demuestra experimentalmente que es prácticamente imposible conseguir dos pruebas exactamente equivalentes. Por eso se tiene que proceder a la igualación y graduación de formas paralelas. En nuestro caso hemos seguido el método equipercentil de igualación y el método lineal de calibrado, presentados ambos por Angoff (1971). El objetivo de este proceso es poder establecer comparaciones entre distintos niveles educativos por una parte, y por otra poder utilizar las puntuaciones convertidas en cálculos posteriores.

El resumen de datos estadísticos que se ofrece permite valorar la bondad de la POLC.

## INTERPRETACIÓN REFERIDA AL CRITERIO Y DIAGNÓSTICO ANALÍTICO.

La evaluación referida al criterio es aquella en que el rendimiento del alumno se compara con los objetivos que se han logrado (criterio), sin tener en consi-

deración los resultados de otros individuos de su grupo. La POLC establece para cada sujeto su «status» respecto a los objetivos de lengua catalana formulados previamente a la elaboración de la prueba. En la presentación de este apartado nos hemos inspirado en Cabrera (1984), a la cual remitimos para una exposición más detallada del proceso.

Recordemos que una interpretación referida al criterio puede ser dicotómica (dos niveles: «domina», «no domina» el objetivo) y polítómica (con diversas categorías o niveles de dominio del objetivo). En la POLC se establecen puntuaciones polítómicas con cuatro categorías: 1) «Buen dominio» del objetivo, contesta a más del 75% de los ítems referidos al objetivo determinado; 2) «dominio inseguro», responde correctamente entre el 50% y el 75% de los ítems referidos al objetivo; 3) «no domina» el objetivo, contesta correctamente entre el 25% y el 50%; 4) «desconocimiento» del objetivo, responde correctamente a menos del 25% de los ítems referidos al criterio.

La interpretación referida al criterio posibilita un diagnóstico analítico del dominio de la lengua catalana en cada individuo. Se ofrecen unos impresos para reflejar el perfil resultante de cada sujeto. Puesto que cada una de las tres formas paralelas de la POLC tiene 100 ítems, esto facilita la interpretación general, ya que la puntuación directa coincide con el porcentaje de aciertos de cada alumno. En base a ese diagnóstico analítico se pueden proponer ejercicios de recuperación de los aspectos que no se dominan suficientemente, tanto a nivel individual como grupal.

## CONCLUSIONES

La Prueba Objetiva de Lengua Catalan (POLC), con tres formas paralelas, (Bisquerra. R. 1985), reune las condiciones de alta fiabilidad y validez para medir el rendimiento en lengua catalana. Se ofrecen unos baremos provisionales para una interpretación normativa basada en centiles. Se establece una interpretación referida al criterio, con puntuaciones politomizadas, que permiten un diagnóstico analítico individual y grupal.

## REFERENCIAS BIBLIOGRÁFICAS

- ANGOFF, W.H. (1971): «Scales, Norms and Equivalent Scores», en R.L. Thorndike, *Educational Measurement, American Council on Education*, Washington, D.C., pp. 508-600.
- BISQUERRA, R. (1985): *Prova objectiva de llengua catalana (POLC)*, CEAC, Barcelona.
- BLOOM, B. S. et al. (1975): *Taxonomía de los objetivos de la educación*, (2 tomos). Marfil, Alcoy.

CABRERA, F. (1984): «Estadio inicial de la lectura. Evaluación diagnóstico-analítica del rendimiento lector». *Revista de Investigación educativa*, vol. 2, n.º 3, pp. 19-39.

COMISSION MIXTA MINISTERI D'EDUCACIÓ I CIENCIA-GENERALITAT DE CATALUNYA, (1979 a), *Programació de Llegua i Literatura Catalanes per al BUP*. Servei Central de Publicacions de la Generalitat de Catalunya, Barcelona.

MAÑA, T. y RIBAS, T. (1982), *Proposta de programació de llengua catalana de 5è. a 8è. d'EGB*. Rosa Sensat-Edicions 62, Barcelona.

RODRÍGUEZ DIEGUEZ, J.L. (1980), *Didáctica General. 1 Objetivos y Evaluación*. Cincel-Kapelusz, Madrid.

# **Impulsividad y Rendimiento Académico: Un Modelo de Modificación del Estilo Cognitivo.**

*por*  
*Leonor Buendía*  
*Juan Ruiz*

## **PLANTEAMIENTO DEL PROBLEMA.**

En estudios anteriores realizados en nuestro departamento sobre la relación entre estilo cognitivo reflexividad-impulsividad y lectura, hemos podido constatar como los niños impulsivos cometan de forma significativa más errores en las tareas lectoras evaluadas que sus compañeros reflexivos. No obstante, parece que dicha dificultad se hace extensiva a gran número de tareas escolares. (Palacios, J. 1982, b) afectando de manera sensible al rendimiento global de los alumnos. Messer S.B. (1976) encontró que los sujetos impulsivos tienen un peor rendimiento académico que los reflexivos; Zelniker T. y Jeffrey W. (1979); Kogan N. (1981); Brow R.T. y Wyne M.E. (1984), han manifestado, así mismo, la alta relación entre impulsividad y bajo rendimiento académico.

A partir de estas evidencias, surgen diversos intentos de modificación del estilo cognitivo, con el fin de ayudar a los sujetos impulsivos, dotándolos de estrategias que les permitan un enfrentamiento más reflexivo con las distintas tareas en las que presentan dificultades.

## **Distintos enfoques al problema de la modificación del Estilo Cognitivo.**

Las investigaciones realizadas en los últimos años sobre la modificación de la impulsividad, podemos dividirlas en tres grandes grupos o vertientes en función de los procedimientos utilizados:

### **a. Utilización de modelos.**

El primer procedimiento de modificación del E.C. impulsividad, se basó en la utilización de modelos reflexivos. Es el procedimiento utilizado entre otros por Yando R. y Kagan, J. (1970); Nagle R.J. y Thwaite B.C.

(1979); Medichenbaum D.H. y Goodman, (1971); Denni (1972); Peters R.S. y Davies K. (1981) y Huhn (1981).

- b. *Alargamiento de los tiempos de latencia.* Desde esta perspectiva se ha intentado la modificación del E.C., forzando la utilización de tiempos de latencia más largos, por parte de los sujetos impulsivos. En este modelo se encuadran los trabajos de: Kagan, J., Pearson L. y Welch L. (1976); Heider E.R. (1971); Pratt M.W. y Wickens G. (1973).

La prolongación de los tiempos de latencia, por sí sola, no ha llevado aparejada una disminución de los errores, por lo que este tipo de procedimientos resultan de poca utilidad.

- c. *Utilización de estrategias cognitivas.*

En esta línea encontramos las investigaciones más prometedoras, puesto que se logró no sólo el pretendido aumento de latencia, sino lo que es más importante la disminución de los errores. Egeland, (1974); Zelniker T. y Jeffrey, W. (1976); Meichenbaum D.H. (1968, 1971, 1977); Clarke, North y Chapel, (1980); Huhn R.H. (1981); Margolis H. (1982); Lin A. Blackman L., Klark H. y Gordon R. (1983); Learner R.W. y Richman C.L. (1984).

Estos trabajos reposan en la cobicción comprobada de que los niños impulsivos manifiestan significativamente menos control verbal de la conducta motora y a la vez usan el lenguaje interno de forma menos instrumental que los reflexivos. En estas investigaciones tienen gran influencia los trabajos de Luria A.R. (1980) y Vygotski L.S. (1980).

Meichenbaum D.L. (1971), con preescolares de 5 años, comprobó que los sujetos reflexivos usan su habla privada de una forma autodirigida más madura y más instrumental que los impulsivos.

A partir de aquí sugiere enseñar a los niños impulsivos a hablarse a sí mismos de forma auto-reguladora; esto daría lugar al cambio conductual y a su vez, al autocontrol. Se trata, en definitiva, de que el niño impulsivo utilice lo que Luria A.R. (1982), denominó «Función Directiva del Habla». Es este el modelo seguido en la presente investigación.

## PLANTEAMIENTO DE LAS HIPÓTESIS.

El objetivo de nuestra investigación es comprobar la eficacia del procedimiento de autoinstrucción propuesto por Meichenbaum D.H. (1971-1977) para la modificación del estilo cognitivo-impulsividad. En función de este objetivo nos planteamos las siguientes hipótesis:

- a) La utilización del procedimiento de autoinstrucción por parte de niños impulsivos no modifica sus tiempos de latencia.
- b) La utilización del procedimiento de autoinstrucción por parte de niños

impulsivos no modifica los errores. El nivel de significación propuesto para la aceptación o rechazo de nuestras hipótesis es del .05.

## MÉTODO

### Sujetos.

La muestra está constituida por 69 niños de primer nivel de E.G.B., correspondientes a dos clases del Colegio de Prácticas de Magisterio de Jaén. De estos 69 alumnos, se seleccionaron sólo los catalogados como impulsivos por el test M.F.F. 20 de Cairns S. y Cammock T. (1978), pero como lo que nos interesa es estudiar la modificación de la impulsividad, cuando esta lleva aparejada problemas de eficiencia, hemos seleccionado para nuestra investigación los niños clasificados como impulsivos y a la vez no eficientes. Estos eran un total de 19 sujetos.

Se han distribuido totalmente al azar en el grupo experimental y grupo de control, quedando cada grupo formado por 9 sujetos. Uno ha quedado sin asignar, con el fin de que los dos grupos tuvieran el mismo «n».

### Instrumentos.

Para la medición de la variable impulsividad se ha utilizado, tanto en el pretest como en el postest, el M.F.F. 20 de Cairns S. y Cammock T. (1978 y 1982). Este nos indica los errores y los tiempos de latencia. El pase fue individual, por el mismo examinador y con las mismas condiciones ambientales para todos los alumnos. Para obtener los valores de impulsividad y eficiencia se utilizaron los índices I y E de Salkind y Wright que proceden de las latencias brutas y de las puntuaciones de error transformadas ambas en puntuaciones «z» a través de las siguientes fórmulas:

$$I_i = Ze_i - Zl_i$$

$$E_i = Ze_i + Zl_i$$

De donde  $ei$  se refiere a las puntuaciones de error y  $li$  a las puntuaciones de latencia del mismo sujeto.

La variable experimental, ha consistido en la utilización del procedimiento de autoinstrucción en la realización de las tareas propuestas; (Meichenbaum D. H. 1971).

### Procedimiento.

Los alumnos impulsivos fueron asignados aleatoriamente a los dos grupos: experimental y control.

Todos están agrupados en dos clases, ubicadas en la misma zona del Colegio, disponiendo de una tercera aula de uso común en la que se han realizado el pretest y postest y sesiones experimentales. El uso de esta aula es diario para gran tipo de actividades y experiencias, por lo que no ha resultado «extraña» a los alumnos.

Las sesiones experimentales se han desarrollado a lo largo de cuatro semanas. En cada una se realizaron cinco sesiones de 35 a 40 minutos. Siempre tuvieron lugar por la mañana.

Mientras los alumnos del grupo experimental recibían el tratamiento utilizando la técnica de autoinstrucción en la realización de las tareas propuestas, los alumnos del grupo de control realizaban las mismas actividades en otra clase, sin utilizar dicho procedimiento.

Las tareas realizadas corresponden a los contenidos del bloque temático n.º 5 del área de Matemáticas del Ciclo Inicial de E.G.B. -Aspectos topológicos y aspectos geométricos.

## RESULTADOS

Se realizó un análisis de covarianza con los datos de errores en el test y otro con los tiempos de latencia.

En ambos casos se cumplió el principio de homocedastidad, comprobado mediante un contraste de hipótesis. El Cociente de las varianzas fue de 1,72 y 1,94 para errores y latencias respectivamente. En ningún caso significativo al .05.

### *Errores en M.F.F. -20*

$\bar{X}$  del valor x nivel I = 41.88889.

$\bar{X}$  del valor y nivel I = 16.55556.

$\bar{X}$  del valor x nivel 2 = 40.77778.

$\bar{X}$  del valor y nivel 2 = 36. 11111.

Estimación del coeficiente de regresión  $\beta = .6431996$

Valor F para el coeficiente  $\beta = 19.23068$ .

Error de muestreo = 1466724.

Estimación de la diferencia de Efectos.

AEXP - ACONT. = -20. 27022.

F = 61.85918.

$\alpha = .05$ ; F (1,15) = 4,54.

*Tiempos de Latencia.*

$\bar{X}$  del valor x nivel 1 = 7.11111  
 $\bar{X}$  del valor y nivel 1 = 16.44445.  
 $\bar{X}$  del valor x nivel 2 = 7.77778.  
 $\bar{X}$  del valor y nivel 2 = 9.2222.

Estimación del coeficiente de regresión  $\beta = .7392856$ .

Valor F para el coeficiente  $\beta = 10.43554$ .

Error de muestreo = .2288521.

Estimación de la diferencia de Efectos.

AEXP - ACONT = 7.71508.

F = 22.65416.

$\alpha = .05$ ; F (1,15) = 4,54.

## DISCUSIÓN

Los resultados obtenidos en el análisis de covarianza indican el efecto de la variable x. (tratamiento con auto-instrucción) sobre el grupo experimental.

Los resultados son altamente significativos, tanto para la variable errores como para los tiempos de latencia en el M.F.F. 20.

Esta disminución de errores nos parece muy importante, pues si bien es cierto que un estilo cognitivo no debe primar sobre otro, en gran número de tareas de la vida, también lo es, que las exigencias de nuestro sistema educativo encuadran mejor con las estrategias preferidas por los reflexivos. Un porcentaje muy elevado del currículum está integrado por tareas que requieren aproximación de detalle, lo que supone una clara desventaja para los alumnos que se conducen de modo impulsivo. En este sentido creemos que la eficacia demostrada por el procedimiento de autoinstrucciones, que hemos investigado, bien puede ser tenido en cuenta en los primeros niveles escolares, a fin de proporcionar a los niños, ya desde los primeros años, las destrezas cognitivas y metacognitivas necesarias, para aplicarlas a toda una variedad de tareas escolares que le proporcionará un mayor Rendimiento en los primeros cursos académicos, con toda la importancia que esto conlleva para su éxito futuro.

## REFERENCIAS BIBLIOGRÁFICAS

- BROWN, R.T. y WYNE, M.E. (1984), «Attentional characteristics and teachers rating in hyperactive reading disabled, and normal boys». *Journal of clinical child psychology*. Spr. Vol 13 (I) 38-43.
- CAIRNS, E., y CAMMOCK, T. (1978), «Development of a more reliable version of the Matching Familiar Figures Test». *Developmental Psychology*, 5, 555-560.
- CAIRNS, E. y CAMMOCK, T. (1982), «Preliminary norms for the MFF 20. Documento inédito.
- DENNEY, D.R. (1972), «Modeling effects upon conceptual style and cognitive tempo». *Child Development*, 43, 105-119.
- HEIDER, E.R. (1971), «Information processing and the modification of an impulsive conceptual tempo. *Child Development*, 42, 1276-1281.
- HUHN, R.H. (1981), «A meta-cognitive approach for teaching cognitive strategies to facilitate learning. RSM2P». *Paper presented at the American Reading Forum annual meeting*. Sarasota, Fla., 7 pp. (ED 211 946).
- KAGAN, J.; PEARSON, L. y WELCH, L. (1966), «Modifiability of an impulsive tempo». *Journal of Educational Psychology*, 57, 359-365.
- KOGAN, N. (1981), «*La psicología en la práctica educativa*. Trillas. México.
- LEARNER, K.M. y RICHMAN, C.L. (1984), «The effect of modifying the cognitive tempo of reading disabled children of reading comprehension». *Contemporary Educational Psychology*, Apr. Vol. 9 (2) 122-134.
- LIN, A., BLACKMAN, L. CLARK, H. y GORDON, R. (1983), «For generalization of visual analogies strategies by impulsive and reflective EMR students». *American Journal of mental Deficiency*, Nov. Vol. 88 (3) 297-306.
- LOGAN, J.W. (1983), «Cognitive style and reading». *Reading Teacher*, Mar, Vol. 36 (7) 704-707.
- LURIA, A.R. (1980), «*Conciencia y lenguaje*». Pablo del Río editor.
- LURIA, A.R. (1982), «Desarrollo y disolución de la función directiva del habla». En *Lenguaje y Psiquiatría*. Ed. Fundamentos.
- MARGOLIS, H. (1982), «Conceptual tempo as moderator variable in predicting first grade achievement test scores». *Journal of School Psychology*, Win, vol 20 (4) 313-322.
- MEICHENBAUM, D.H. (1968), «Modification of classroom behavior of institutionalized female adolescent offenders». *Behavior Research Therapy*. 6, 343-353.
- MEICHENBAUM, D.H. (1968), «Reflection-impulsivity and verbal control of motor behavior». *Child Development*, 40, 785-797.
- MEICHENBAUM, D.M. y GOODMAN, J. (1971) «Training impulsive children to talk to themselves: a means of developing self-control». *Journal of Abnormal Psychology*, 77, 115-126.

- MEICHENBAUM, D.H. (1971, b), «The nature and modification of impulsive children: training impulsive children to talk to themselves». Manuscrito inédito (adaptación de un trabajo presentado para la conferencia de 1971 de la SRCD celebrada en Minneapolis, Minnesota).
- MEICHENBAUM, D.H. (1974) «Self-instructional strategy training» *Human Development*, 273-280.
- MEICHENBAUM, D.H. (1977) «Cognitive-behavior modification.» *An integrated approach*. Plenum press, New York.
- MEICHENBAUM, D.H. (1980), «Toward a conceptual model for the treatment of test anxiety. En R.G. Sarason: *Test anxiety*. Erlbaum.
- MESER, S.B. (1976), «Reflection-impulsivity: a review.» *Psychological Bulletin*. 83, 1026-1052.
- NAGLE, R.J. y THAWAITE, B.C. (1979), «Modeling effects on impulsivity with learning disabled children». *Journal of Learning Disabilities*, May, Vol. 12 (5) 331-336.
- PALACIOS, J. (1982), «Implicaciones educativas de los estilos cognitivos». *Rev. Infancia y Aprendizaje*. n.º 18, 84-106.
- PETERS, R.S. y DAVIES, K. (1981), «Effects of self-instructional training on cognitive impulsivity of mentally retarded adolescents». *American Journal of Mental Deficiency*, Jan, Vol. 85 (4) 377-382.
- PRATT, M.W. y WICKENS, G. (1983), «Checking it out: cognitive style, context, and problem type in children's monitoring of text comprehension». *Journal of Educational Psychology*, Oct. Vol. 75 (5) 716-726.
- VYGOTSKY, L.S. (1980), «*Pensamiento y lenguaje*». La Pléyade, Buenos Aires.
- YANDO, R. Y KAGAN, J. (1970), «The Effects of task complexity on reflection-impulsivity». «*Cognitive Psychology*», I, 192-200.
- ZELNIKER, T. y JEFFREY, W. (1976), «Reflective and impulsive children: strategies of information processing differences in problem solving». *Monographs of the Society for Research in Child Development*, 41. (n.º 5 completo).
- ZELNIKER, T. y JEFFREY, W. (1979), «Attention and cognitive style in child». En G.A. Hace. y M. Lewis: *Attention and cognitive development*, Plenum Press, 275-296, New York.
- ZELNIKER, T., JEFFREY, W., AULT, R. y PEARSON, J. (1972), «Analysis and modification of search strategies of impulsive and reflective children on the Matching Familiar Figures Test». *Child Development*, 43, 231-235.

# Aplicación del Meta-Análisis al estudio de algunos factores del rendimiento escolar: una investigación en Galicia

por  
*José Cajide y Xose Rubal*

Cuando se intenta explicar el rendimiento escolar es preciso recurrir, entre otras, a las características del alumno en los ámbitos personal, familiar, las condiciones escolares, y las actitudes, (Binimianov, I y Glasman, N. 1983). De ellas nos ocuparemos en este trabajo. No obstante, queremos advertir que no constituye este estudio un análisis exhaustivo de dichos factores; más bien, tratamos de aprovechar los datos de dos investigaciones haciendo una presentación pedagógicamente contextualizada de una de las técnicas del meta-análisis: el cálculo del tamaño del efecto. Somos conscientes de que existen otras vías de análisis alternativas, pero estimamos que puede ser rentable dicha técnica al pretender, desde una perspectiva diferenciada en cuanto a las materias, ofrecer una aproximación a la estimación del peso explicativo que sobre el rendimiento escolar tienen diversas variables habitualmente empleadas como predictores.

## VARIABLES

De los seis grupos de variables del modelo de Binimianov, I. Glasman, N. nosotros hemos incluido en nuestro estudio cuatro. De un lado, como variable dependiente, estaría el rendimiento en las disciplinas: lengua castellana, gallega, matemáticas y sociales; es decir, una parte representativa de los resultados cognitivos en las áreas fundamentales del ámbito escolar, áreas de lengua, matemáticas y sociales. El criterio utilizado para la constatación del rendimiento, fue la calificación escolar. De otro, tendríamos tres grupos de variables, las características generales de base del alumno, representados por el sexo (condición individual), la de la clase social (condición socio-familiar), y la lengua materna del alumno (condición socio-lingüística).

## INSTRUMENTOS DE MEDIDA

Los datos de sexo y lengua materna proceden de autoinformes de los alumnos; la clase social se derivó a partir de la profesión y estudio de los padres; las condiciones de la institución escolar se obtuvieron del carácter público o privado de los centros; las actitudes respecto a las materias se midieron mediante un grupo de ocho adjetivos bipolares (diferencial semántico), que responden a una dimensión factorialmente consistente en la evaluación de las cuatro materias, con fiabilidades comprendidas entre los valores alfa .85 y .89. Las llamamos sintéticamente «actitud».

## MUESTRA

En los estudios de meta-análisis la muestra la constituyen normalmente un conjunto de investigaciones relativas al mismo problema. En nuestro caso, en el ámbito gallego, no era posible por carecer de investigaciones; de ahí que partimos de dos estudios en los que se consideran cinco sub-problemas relativos a la explicación del rendimiento, similares en estructura y resolución; pero diferentes en razón del referente central del análisis: la variable dependiente en cuestión, el rendimiento en cada una de las materias citadas.

Los sujetos son en un caso alumnos de 6.<sup>º</sup> y 8.<sup>º</sup> de E.G.B., (N = 301), y en el otro, alumnos de 2.<sup>º</sup> de B.U.P., (N = 429).

## MÉTODO

El análisis de los datos fue realizado a través del análisis factorial de varianza, con el ánimo de separar los efectos de cada una de las variables independientes sobre las dependientes, así como, las interacciones. Posteriormente, se calculó la magnitud del tamaño del efecto de todas las variables independientes sobre cada una de las variables dependientes. Para obtener el tamaño del efecto nos basamos en la fórmula de (Glass, G.V. et alt., 1981, p. 119):

$$d_i = \frac{\bar{X}_i^E - \bar{X}_i^C}{\sqrt{\frac{SC_A + SC_B + SC_C + SC_{AB} + SC_R}{gl_A + gl_B + gl_C + gl_{AB} + gl_R}}}$$

donde:

$d_i$  es el tamaño del efecto;  $\bar{X}_i^E - \bar{X}_i^C$  son las medias respectivas de cada una de las variables independientes; las sumas de cuadrados (SC), divididas por los grados de libertad (gl), corresponden al cálculo de la desviación standard, por la cual

ha de partirse cada diferencia de medias para el cálculo del tamaño del efecto. Por ejemplo, para el cálculo del tamaño del efecto del rendimiento en lengua castellana a partir de la variable clase social, la media de la clase social alta fue 3,56 y la de la clase social baja, 3,13. Las sumas de cuadrados y grados de libertad necesarios son los del tipo de centro, lengua materna, sexo, actitud, la interacción, que nosotros añadimos a la residual por no ser significativa, y la residual, cuyos valores figuran bajo la raíz:

$$d_i = \frac{3,56 - 3,13}{\sqrt{\frac{1,806 + 2,057 + 4,458 + 3,291 + 383,955}{1 + 1 + 1 + 1 + 273}}} = .33$$

## RESULTADOS

Tanto los análisis factoriales de varianza, como los cálculos del tamaño del efecto, nos vemos obligados a reducirlos a razones de F e índices el tamaño del efecto, dadas las exigencias de la síntesis. Por ello, las tablas 1-4 presentan sólo estos resultados:

**TABLA I**  
Valores F y tamaño del efecto para el rendimiento  
en castellano

	Clase Social	Tipo Centro	Lengua materna	Sexo	Actitud
F	7.667*	1.284	1.463	3.170	2.340
d <sub>i</sub>	.33	.24	.32	.36	.34

\* Los valores F seguidos de asterisco indican que son significativos al .05.

**TABLA II**  
Valores F y tamaños del efecto para el rendimiento en gallego  
(muestra de E.G.B.)

	Clase Social	Tipo Centro	Lengua materna	Sexo	Actitud
F	9.179*	1.168	.143	3.255	4.310*
d <sub>i</sub>	.38	.11	.07	.21	.31

**TABLA III**  
Valores F y tamaños del efecto para el rendimiento en matemáticas

	Clase Social	Tipo Centro	Lengua materna	Sexo	Actitud
F	1.966	1.432	4.371*	1.121	3.894*
d <sub>i</sub>	.39	.32	.49	.11	.45

**TABLA IV**  
Valores F y tamaños del efecto para el rendimiento en sociales

	<u>Clase Social</u>	<u>Tipo Centro</u>	<u>Lengua materna</u>	<u>Sexo</u>	<u>Actitud</u>
F	8.786*	.04	4.229*	2.807	7.805*
d <sub>i</sub>	.36	.10	.39	.37	.49

**TABLA V**  
Valores de F y tamaños del efecto para el rendimiento en gallego  
(muestra Bachillerato)

	<u>Clase Social</u>	<u>Tipo Centro</u>	<u>Lengua materna</u>	<u>Sexo</u>	<u>Actitud</u>
F	1.691	Se carece de datos	.321	1.040	10.686*
d <sub>i</sub>	.21		.17	.19	.55

**TABLA VI**  
Tamaños del efecto de cada una de las variables independientes sobre  
todas las variables dependientes combinadas.

Variables	n.º de tamaños del efecto	d*	Desviación standard
Clase social	4	.37	.002
Tipo de centro	4	.19	.09
Lengua materna	4	.32	.15
Sexo	4	.24	.09
Actitud	4	.40	.07

\*d se estimó por  $\sum_{i=1}^J d_i/J$ , es decir, un simple promedio aritmético.

## INTERPRETACIÓN DE LOS RESULTADOS

El análisis de las tablas precedentes nos indica que las variables cuyos efectos sobre el rendimiento en las distintas materias, resultan significativas un mayor número de veces, son: la clase social y la actitud. La primera lo sería en relación con el rendimiento en castellano, gallego y sociales; mientras que la segunda lo es en gallego, matemáticas y sociales. La variable lengua materna lo sería curiosamente, sólo en relación con las matemáticas y sociales, y no lo serían en ningún caso, las variables tipo de centro y sexo. Igualmente apreciamos como la variación en lengua materna no daría lugar a efectos significativos en el rendimiento en materias del área lingüística, y sí en las áreas de matemáticas y sociales. A su vez, los efectos del tipo de centro y sexo, presentan los valores más

bajos. Debe resaltarse que el efecto diferencial de las variables independientes se produce en todas las materias de forma uniformemente favorable a las categorías: clase social alta, centro privado, lengua materna castellana, (incluso para el rendimiento en gallego), y actitud favorable superior a la neutra; mientras que en el caso del sexo, variaría. Se produciría un efecto diferencial a favor de las mujeres en las materias de gallego y sociales; y a favor de los hombres en lengua castellana y matemáticas.

Si consideramos la importancia explicativa de las distintas variables desde la óptica de la magnitud del tamaño del efecto, éste puede ser interpretado como la «puntuación zeta» de realización para unos sujetos promedio en el grupo experimental relacionado con el desempeño promedio de los sujetos en el grupo de control, (Glass, G.V. et alt, p. 103). En nuestras investigaciones, los tamaños del efecto correspondientes a las variables clase social y actitud, son, por lo general, los de mayor magnitud. En la tabla n.º 6 se constata que las respectivas medias combinadas correspondientes a dichas variables son las más altas y las más fiables (menor dispersión). Por último, digamos que nuestros datos replican algunas de las hipótesis verificadas por (Pérez Serrano, G. 1981, pp. 250-255), a excepción del rendimiento según el tipo de centro, dado que en nuestro caso, es superior el rendimiento en los privados, sin que las diferencias sean significativas. Igualmente, habría que resaltar la importancia del componente actitudinal con respecto al rendimiento, puesto de manifiesto por Schibeci, R.A. en 1984, y corroborada por la evidencia empírica de nuestro trabajo.

#### REFERENCIAS BIBLIOGRÁFICAS

- BINIAMINOV, I. & GLASMAN, N.S. (1983): «School Determinants of Student Achievement in Secondary Education». *American Educational Research Journal*, 20, 2, 251-268.
- GLASS, G. V. et alt. (1981): *Meta-Analysis in Social Research*. SAGE Publications, Beverly Hills.
- PÉREZ SERRANO, G. (1981): *Origen social y rendimiento escolar*. Centro de investigaciones sociológicas. Madrid.
- ROSENTHAL, R. (1984): *Meta-analytic Procedures for Social Research*. SAGE Publications, Beverly Hills.
- SCHIBECI, R. A. (1984): «Students Teachers and the Assessment of Attitude of School». *The Australian Journal of Education*, 28, 1, 17-24.

## Modelos interaccionistas. Aportaciones y limitaciones

por  
*M.ª Pilar Colás*

La preocupación en la última década por la enseñanza individualizada ha impulsado el desarrollo de estudios en los que se trata de adecuar los tratamientos educativos a las peculiaridades individuales de los alumnos.

La conjugación de variables instructivas y discípulas requiere diseños de investigación adecuados. Los diseños ATI (Interacción Tratamiento y Aptitud) permiten cumplir este propósito. Cronbach, L. (1981 a) fue el primero que, a la vista de la parcialidad de las dos líneas de investigación, la correlacional y la experimental para abordar los problemas de interacción, aboga por una línea integradora a la que denominará ATI. En estos diseños se consideran los tratamientos y aptitudes simultáneamente. El tratamiento estadístico se formaliza en un análisis de varianza que determina la significación de ATI.

El término «aptitud» podría ser definido como alguna característica de la persona que afecte a su respuesta al tratamiento (Cronbach, L. 1981 b). Estas características pueden referirse al dominio cognitivo o a variables de personalidad.

Los diseños ATI pretenden detectar el grado y forma de la interacción entre dos tipos de variables. La utilización de diseños ATI en el campo educativo exige una previa selección de tratamientos que puedan adaptarse a alumnos con aptitudes peculiares, y de aptitudes que interactúen con aspectos modificables de los tratamientos.

La determinación de aptitudes y tratamientos educativos representativos puede realizarse a partir de modelos teóricos previos. Snow, R. (1970) y Salomon, G. (1972) proponen algunos modelos para conceptualizar y formalizar los diseños ATI:

a) *Modelo Remediador*.— La base de la que parte es que la instrucción puede remediar las deficiencias del aprendizaje. En este modelo se hace necesario un conocimiento secuencial y jerárquico de cómo se aprende y de la función o misión de la enseñanza en cada fase de este proceso. Este planteamiento está estrechamente ligado al pensamiento de Gagné y Ausubel entre otros.

Las variables instructivas y aptitudes que aporta este modelo están íntimamente relacionadas con procesos de aprendizaje, atendiendo a las cuestiones de qué es lo que se debe aprender y en qué orden para conseguir un máximo de transferencia. Las aptitudes que se contemplan en este modelo se refieren a habilidades altamente específicas. Los tratamientos recogen estrategias didácticas acordes a los distintos niveles de habilidades específicas.

b) *Modelo Compensatorio*.- En este modelo los tratamientos que interactúan con las aptitudes compensan las deficiencias del aprendiz, proveyendo un modo de presentación que el alumno no puede proponer por sí mismo. Las variables de instrucción presentan aspectos generales de la instrucción, y se refieren preferentemente a formas de presentación de información. Las aptitudes poseen un mayor grado de generalidad que en el caso anterior, a la vez que son permanentes y no se modifican como en el caso anterior.

c) *Modelo Preferencial*.- Los tratamientos se conjugan de acuerdo con el estilo o estrategia de procesamiento preferente del discente. El tratamiento instrucción 1 es diseñado en función de la aptitud con la que el aprendiz está mejor equipado. Este modelo se ocupa de capacidades bien desarrolladas en oposición a los anteriores que se ocupan de deficiencias.

Cada uno de estos modelos propicia hipótesis de trabajo con particulares aptitudes y tratamientos. Los diseños ATI pueden asumir diferentes modelos.

Las aportaciones empíricas de los diseños ATI, llevados a cabo en el campo de la enseñanza pueden agruparse alrededor de cuatro contenidos temáticos, siguiendo la propuesta de Cronbach, L. y Snow, R (1977):

- a) Estudios que relacionan aptitudes generales con tratamientos referidos a características de los estímulos.
- b) Relaciones entre aptitudes generales discentes y, composición de la estructura didáctica.
- c) Interacción entre tratamientos concernientes al procesamiento de la información y aptitudes generales.
- d) Métodos de enseñanza, y estilos cognitivos.

Los trabajos que se incluyen que se incluyen en el primer apartado pueden agruparse dentro de los denominados modelos compensatorios. Tratan de averiguar la representatividad de los estímulos visuales frente a los verbales en contenidos espaciales en aquellos sujetos que poseen escasas habilidades espaciales. En algunos de estos trabajos (Salomon, G. 1974) se confirma la interacción entre variables de tratamientos y variables de aptitud.

En la compoción de la estructura didáctica se tratan las formas de elaboración de textos, las preguntas y los organizadores previos, entre otras. Con esta temática se han elaborado diseños pertenecientes a modelos compensatorios y preferenciales. Algunas aportaciones sobre estos contenidos (Ausubel D. y Fitzgerald, D. 1962, Shavelson, R. y Col, 1974) y apuntan el interés de estas variables de enseñanza para relacionarse con características cognitivas de los sujetos.

Los tratamientos referidos a formas de procesamiento de la información se han concretado en los métodos de enseñanza inductivo y deductivo. Las aptitudes que se han interaccinado con estos tratamientos han sido de carácter general y específico, considerándose aptitudes matemáticas y verbales. A pesar de que en casi todos los estudios se manifiesta una clara correlación entre aptitud y rendimiento, no se obtienen interacciones positivas.

Los estilos cognitivos de los alumnos han tenido una especial consideración en los actuales trabajos de investigación relacionándose con los métodos de enseñanza. Los resultados no han sido unánimes. Mientras en algunos trabajos (Grieve, T. y Davis, J. 1971) la interacción entre estilo cognitivo Dependencia de Campo y método expositivo y descubrimiento resulta positiva, otros autores (Okeke, F., 1983 y Leeuw, L. 1983) no hallan interacciones significativas entre ambos tipos de variables, siendo en la mayoría de los casos el estilo cognitivo representativo del éxito escolar.

En la actualidad los datos aportados por los diseños ATI no pueden generalizarse debido a las discrepancias de resultados. Esta falta de unanimidad exige una reflexión de índole conceptual y metodológica, tema que indudablemente sobrepasa los objetivos de este trabajo y por tanto no abordamos en ésta exposición.

Otros motivos, aparte del mencionado, hacen difícil la generalización; a) escasa aplicación de diseños ATI a contenidos educativos, b) diferente nivel de amplitud de las variables tratadas y c) la no utilización de modelos teóricos en la selección de variables experimentales y en la interpretación de resultados, etc.

Los diseños ATI, como cualquier otro tipo de metodología de investigación, no está exenta de errores y limitaciones que es preciso tomar en consideración. Uno de los puntos débiles que se señalan en estos diseños (Tobias, S. 1976) es la aceptación de la estabilidad de las aptitudes discentes, cuestión que no se admite desde la perspectiva de la psicología cognitiva. Desde el punto de vista metodológico hay que reseñar la dificultad de controlar las numerosas variables que interactúan en el experimento y que pueden tener incidencia determinante en los resultados.

A pesar de estas limitaciones, la información derivada de modelos interaccionistas puede tener una repercusión muy directa en el tratamiento del rendimiento escolar, pudiéndose aplicar a cuestiones relacionadas con el fracaso escolar, la educación compensatoria y la potenciación de habilidades específicas.

## REFERENCIAS BIBLIOGRÁFICAS

- AUSUBEL, D. y FITZGERALD, D. (1962), «Organizar, general background and antecedent learning variables in sequential verbal learning», *Journal of Educational Psychology*, pp. 243-249.
- CRONBACH, L. y SNOW, R. (1977), *Aptitudes and Instructional Methods*, Irvington Publishers, Inc.
- CRONBACH, L. (1981 a), «Las dos disciplinas de la disciplina científica», en Alvira et al. *Los dos métodos de las Ciencias Sociales*, Centro de Investigaciones Sociales.
- CRONBACH, L. (1981 b), «Más allá de las dos disciplinas de la psicología científica», en Alvira et al. op. cit., p. 254.
- GRIEVE, T. y DAVIS, J. (1971), «The relationship of cognitive style and method of instruction to performance in ninth grade geography», *Journal of Educational Research*, 66, pp. 137-111.
- LEEUW, L. (1983), «Teaching problem solving: an ATI study of the effects of Teaching Algorithmic and Heuristic solution methods», *Instructional Science*, 12, pp. 1-48.
- OKEKE, F. (1982), «Cognitive styles, Instructional Strategies, and Academic Performance», *The Journal of Experimental Education*, v. 5, 1, pp. 31-37.
- ROTHKOPH, E. (1972), «Variable adjunct question schedules, interpersonal interaction, and incidental learning from written material», *Journal of Educational Psychology*, 66, pp. 87-92.
- SALOMON, G. (1972), «Heuristic models for the generation of aptitude. Treatment Interaction Hypotheses», *Review of Educational Research*, 42, pp. 327-343.
- SALOMON, G. (1974), «Internalization of filmic operations in relation to individual differences», *Journal of Educational Psychology*, 66, pp. 499-511.
- SHAVELSON, R. y Co. (1974), «The effects of position and type of question on learning from prose material: interaction of treatment with individual differences», *Journal of Educational Psychology*, 66, pp. 40-48.
- SNOW, R. (1970), «Research on media and aptitudes», *Bulletin of the Indiana University School of Education*, cit. por Cronbach y Snow (1977).
- TOBIAS, S. (1976), «Achievement treatment interaction», *Review of Educational Research*, 46, pp. 61-74.

## **Aplicación del análisis de preferencias en el estudio sobre el rendimiento académico**

*por  
Norberto Corral*

### **INTRODUCCIÓN**

Un objetivo común a los estudios sobre evaluación del rendimiento académico es determinar aquellos factores que influyen sobre él, estudiando la relación entre las características sociales, pedagógicas y sicológicas del alumno y su rendimiento.

En aquellos casos en que podemos conocer directamente del alumno la importancia que dichos factores tienen para él, se plantea la posibilidad de trabajar en forma inversa a la anterior. Es decir determinar primero los diferentes grupos de opinión que sobre dichos factores aparecen en la población y estudiar posteriormente las características individuales de los sujetos pertenecientes a cada grupo. Es en la formación de estos grupos de opinión donde podemos utilizar el análisis de preferencias.

Por tanto la información disponible en un estudio de este segundo tipo podemos clasificarla en:

- a) Información directa sobre la importancia de los factores.
- b) Información sobre los individuos.

El modelo de preferencias de análisis interno ANAPREF, desarrollado por Jacquet-Lagreze sirve como referencia en la realización del trabajo, si bien el criterio utilizado en la formación de los grupos de opinión difiere del anterior, surgiendo a raíz del estudio realizado sobre un colectivo de maestros con antigüedad entre uno y tres años (como funcionarios) que valoraron comparativamente los principales problemas que se planteaban en la escuela.

## PRESENTACIÓN DEL PROBLEMA

Consideramos un conjunto de  $N$  individuos, a cada uno de los cuales se le pide que manifieste sus preferencias sobre  $M$  objetos; los resultados de estas preferencias los representamos mediante una matriz de término general  $a_{ijk}$  y de dimensión  $N \times M$ .

Los términos  $a_{ijk}$  y  $a_{ikj}$  representan, respectivamente, las preferencias del individuo  $i$  de  $j$  sobre  $k$  y  $k$  sobre  $j$ , pudiendo tomar los siguientes valores:

- a.- Si  $j$  es preferido a  $k$  por  $i$   $a_{ijk} = 1$  y  $a_{ikj} = 0$ .
- b.- Si  $j$  y  $k$  tienen igual preferencia  $a_{ijk} = 1/2$  y  $a_{ikj} = 1/2$
- c.- Si  $j$  y  $k$  son incomparables para  $i$   $a_{ijk} = 0$  y  $a_{ikj} = 0$ .

Aplicando el modelo diseñado por Jacquet-Lagreze, representaremos la predilección de  $j$  sobre  $k$  en el grupo que estamos estudiando por:

$A(j,k) = \sum_{i=1}^N a_{ijk}$ . Una propiedad interesante para  $A(j,k)$  y  $A(k,j)$  es la siguiente:

### Propiedad 1:

\*  $\leq A(j,k) + A(k,j) \leq N$ . En efecto:

$$A(j,k) + A(k,j) = \sum_{i=1}^N (a_{ijk} + a_{ikj}) \leq \sum_{i=1}^N 1 = N$$

### Propiedad 2

$$0 \leq \sum_H (A(j,k) + A(k,j)) \leq \sum_H N = N \cdot C_M^2 = \frac{N \cdot M \cdot (M-1)}{2};$$

Siendo  $H =$  conjunto de parejas  $(j,k)$  con  $j \neq k$      $j = 1 \dots M$   
 $K = 1 \dots M$

### Definición:

Se define la *intensidad de opinión* como el cociente:

$$\frac{\sum_H (A(j,k) + A(k,j))}{N \cdot M \cdot (M-1)/2}$$

que se interpretará como el reflejo del conocimiento que sobre los objetos tienen los individuos del grupo o índice de no respuestas.

Nuestro interés, a partir de ahora, podemos centrarlo en encontrar subgrupos del grupo total, que tengan parecida opinión sobre los objetos. Estos subgrupos quedarán caracterizados mediante órdenes globales sobre los objetos. Para ello definimos:

**Definición:**

La preferencia del objeto  $j$  sobre los demás será

$$B^1(j) = \sum_{k \neq j} (A(j,k) - A(k,j))$$

cumpliendo la propiedad  $-N.(M-1) \leq B^1(j) \leq N.(M-1)$

Parece razonable por tanto, si pretendemos formar grupos de acuerdo a las preferencias de los sujetos, elegir como primer objeto aquel que posea el  $B^1(j)$  máximo. Una vez hallado este, el problema se reduce a aplicar el mismo procedimiento sobre los restantes objetos.

Representamos por  $B^r(j) = \sum_{k \neq j, j_1 \dots j_{r-1}} (A(j,k) - A(k,j))$  la preferencia de  $j$  sobre los objetos, exceptuando  $j_1 \dots j_{r-1}$ , con la propiedad de que  $-N.(M-r) \leq B^r(j) \leq N.(M-r)$

**Definición:**

Definimos  $B^r = \max_j (B^r(j))$

$B^r$  representa la aceptación en términos absolutos del objeto colocado en la  $r$ -ésima posición.

**Definición:**

Llamaremos *cohesión del grupo G* con el orden establecido  $O$  al cociente:

$$C(G,O) = \frac{\sum_{r=1}^{M-r} B^r}{N.(M-1).M/2} \text{ con } 0 \leq C(G,O) \leq 1$$

**Definición 1:**

Definimos la matriz de preferencias del orden  $O$  como:

$$O(j,k) = \begin{cases} 1 & \text{si } j > k \text{ en el orden } O \\ 0 & \text{si } k < j \text{ en el orden } O \end{cases}$$

$$O(j,j) = 0.$$

**Definición 2:**

Se define la distancia de un sujeto al orden  $O$ , como

$$d(a_i, O) = \sum_k |a_{ijk} - o_{jk}|$$

Elegimos una distancia máxima admisible DM para asignar un individuo a un orden de acuerdo con el siguiente criterio:

Si  $d(a_i, O) \leq DM$  el individuo  $i \in G(O)$

Si  $d(a_i, O) > DM$  el individuo  $i \notin G(O)$

Sea  $G_i = \{j / d(a_i, O) \leq DM\}$ . Calculamos en  $G_i$  la nueva tendencia y la distancia de los sujetos a ella, formando un nuevo  $G_i$ , reiterando el proceso hasta obtener un grupo estable y el orden asociado a él. Con los restantes sujetos repetimos el método anterior hasta conseguir grupos estables con intersección vacía y los órdenes asociados a ellos.

Por último se hallan las distancias de cada individuo a los órdenes establecidos, asignándolos a aquel grupo cuya distancia sea menor. Sobre estos nuevos grupos se vuelve a calcular la tendencia de opinión.

#### REFERENCIAS BIBLIOGRÁFICAS

- JACQUET-LAGREZE E. (1971) «Opinions valuees et graphes de preferences». *Mathematiques et Sciences Humaines*, n.º 33.
- BERTIER P., BOUROUCHE J.M. (1971); *Analyse des donnees multidimensionnelles*, París: P.U.F.
- BENZEERI J.P. (1965); «Sur l'analyse des preferences» .I.S.U.P., Polycopie.
- LEBART L., MORINEAU A., FENELON J.P. (1982) «Traitement des donees statistiques». París: Dunod.

NÚMERO DE SUJETOS N = 12  
 NÚMERO DE OBJETOS M = 4  
 DISTANCIA MAXIMA ADMISIBLE DM = 6

#### DATOS ORIGINALES

0	0	1	1	0	0	1	1	0	0	0	2	0	0	2	0
0	1	1	1	0	0	0	1	0	1	0	1	0	0	0	0
0	0	1	0	1	0	1	0	0	0	0	0	1	1	1	0
0	0	0	0	1	0	0	0	1	1	0	0	1	1	1	0
0	0	0	0	1	0	1	0	1	0	0	0	1	1	1	0
0	1	1	1	0	0	0	1	0	1	0	1	0	0	0	0
0	0	1	1	0	0	0	1	0	1	0	2	0	0	2	0
0	0	1	1	1	0	1	1	0	0	0	0	0	0	1	0
0	1	1	1	0	0	0	1	0	1	0	1	0	0	0	0
0	1	0	0	0	0	0	0	1	1	0	1	1	1	0	0
0	1	0	1	0	0	0	1	1	1	0	1	0	0	0	0
0	1	1	0	0	0	2	0	0	2	0	0	0	0	1	0
0	0	0	0	1	0	1	0	1	0	0	0	1	1	1	0
0	1	1	1	0	0	0	1	0	1	0	1	0	0	0	0
0	1	1	1	0	0	0	0	1	0	0	0	0	1	1	0
0	0	1	1	0	0	0	1	1	1	0	1	0	0	0	0
0	0	0	1	1	0	1	1	1	0	0	1	0	0	0	0
0	0	0	0	1	0	0	1	1	1	0	1	1	0	0	0
0	0	0	0	1	0	0	1	1	1	0	1	1	0	0	0

- \* EL NÚMERO 2 REPRESENTA LA PUNTUACION 0.5
- \* LAS FILAS REPRESENTAN LA COMPARACION EN CADA SUJETO DE TODOS LOS OBJETOS . POR EJEMPLO LA FILA DOS EQUIVALE AL ORDEN 1 > 3 > 2 > 4
- \* LOS OBJETOS NO SE COMPARAN CON ELLOS MISMOS . A (1.J.J)= 0

#### GRUPO TOTAL

NUMERO DE SUJETOS = 19  
 COHESION DEL GRUPO =.149122807  
 INTENSIDAD DE LA OPINION DENTRO DEL GRUPO =.964912281  
 ORDEN DE LOS OBJETOS EN EL GRUPO = 3 2 14

#### GRUPO 1

NUMERO DE SUJETOS = 11  
 COHESION DEL GRUPO =.545454545  
 INTENSIDAD DE LA OPINION DENTRO DEL GRUPO =.984848485  
 ORDEN DE LOS OBJETOS EN EL GRUPO = 3 1 2 4

#### GRUPO 2

NUMERO DE SUJETOS = 8  
 COHESION DEL GRUPO =.395833333  
 INTENSIDAD DE LA OPINION DENTRO DEL GRUPO = .9875  
 ORDEN DE LOS OBJETOS EN EL GRUPO = 4 2 1 3

## CLASIFICACION DE LOS SUJETOS POR GRUPOS

SUJETO	GRUPO	DISTANCIA
1	2	6
2	1	2
3	2	0
4	2	4
5	2	2
6	1	2
7	1	4
8	2	4
9	1	2
10	1	4
11	1	0
12	2	5
13	2	2
14	1	2
15	2	6
16	1	2
17	1	4
18	1	4
19	1	1

## DATOS ORDENADOS POR GRUPOS

0	1	1	1	0	0	0	1	0	1	0	1	0	0	0	0	0	0
0	1	1	1	0	0	0	0	1	0	1	0	1	0	0	0	0	0
0	0	1	1	0	0	0	0	1	0	1	0	2	0	0	2	0	0
0	1	1	1	0	0	0	0	1	0	1	0	1	0	0	0	0	0
0	1	0	0	0	0	0	0	0	1	1	0	1	1	1	0	0	0
0	1	0	1	0	0	0	0	1	1	1	0	1	0	0	0	0	0
0	1	0	1	0	0	0	0	1	1	1	0	1	1	1	0	0	0
0	1	1	1	0	0	0	0	1	0	1	0	1	0	0	0	0	0
0	0	1	1	0	0	0	0	1	1	0	1	0	0	0	0	0	0
0	0	0	1	1	0	0	1	1	1	0	0	1	0	0	0	0	0
0	0	0	0	1	0	0	0	1	1	1	0	1	1	0	0	0	0
0	0	0	0	1	0	0	0	1	1	1	0	1	1	0	0	0	0
0	0	0	0	0	1	0	0	0	1	1	0	0	2	0	0	2	0
0	0	1	1	0	0	0	1	1	0	0	0	0	1	1	1	0	0
0	0	1	0	1	0	0	1	0	0	0	0	0	1	1	1	0	0
0	0	0	0	1	0	0	0	1	1	0	0	0	1	1	1	0	0
0	0	0	0	1	0	0	1	0	1	0	0	0	0	1	1	1	0
0	0	1	1	1	0	1	1	0	0	0	0	0	0	0	1	0	0
0	1	1	0	0	0	2	0	0	2	0	0	0	0	0	0	1	0
0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	1	1	0
0	1	1	1	1	0	0	0	0	1	0	0	0	0	0	1	1	0

# Metodología para la valoración de la actividad hospitalaria del pregraduado en cirugía

por  
M. Diez Miralles y otros

## OBJETIVOS DEL ESTUDIO

Se pretende exponer una metodología científica que permite valorar de una forma objetiva la eficacia de un sistema docente de aspecto práctico, utilizando técnicas de Investigación educativa descriptiva y experimental.

Para ello se ha elegido el modelo de enseñanza práctica de la Cirugía que se desarrolla en la Facultad de Medicina de Alicante.

## MATERIAL Y MÉTODOS

Se ha valorado la Actividad Hospitalaria (A. H.) de los estudiantes que acudieron al Servicio de Cirugía General del Hospital General de Elche, durante los cursos 1983-84 y 1984-85. Se han empleado sistemas de valoración:

A) **Estudios de tipo Encuesta**, dirigidos a las siguientes muestras de población:

1- 87 Estudiantes de *cuarto curso de medicina* en los cursos mencionados, que han sido motivo del presente modelo de A. H.

2- 80 Pacientes tratados durante el periodo estudiado en nuestro Servicio.

Estas encuestas se plantearon en forma de cuestionario directo y anónimo, individual, y con preguntas cuya respuesta supone un sistema doble de recogida de información:

– Respuesta cerrada en forma de una elección sobre varias opciones, planteada como de obligada contestación.

– Respuesta abierta planteada al encuestado como de contestación totalmente opcional.

La valoración estadística de los datos, se realizó mediante la  $\chi^2$  de Pearson cuando los resultados estuvieron tabulados en forma de distribución de frecuencias de la variable, y que proporciona información inferencial sobre si la diferencia encontrada es estadísticamente significativa respecto de las frecuencias teóricas esperados. Cuando las frecuencias esperadas están en el intervalo entre 3 y 5 hemos aplicado la correlación de Yates a aquel estadígrafo.

Si los resultados se ofrecen en forma de proporción porcentual se aplicó la aproximación a la normal de cada respuesta a cada pregunta, tras dividir la diferencia proporcional por el error típico.

Para valorar si las respuestas a dos preguntas estaban relacionadas entre sí hemos construido tablas de doble entrada que hemos valorado mediante la aplicación de la  $\chi^2$ .

**B) Estudios de tipo Evaluación Formativa.** Han tenido una doble finalidad formativa y experimental, pero sin repercusión certificativa. Se diseñaron según el sistema de Preguntas de Elección Múltiple con una sola respuesta válida. Estos Tests han constado de 20 preguntas, referente a nociones o conocimientos adquiridos en el periodo de A. H., de nueva creación en su mayoría, puesto que no abundan en la literatura con la finalidad de evaluar conocimientos de prácticas de Cirugía. Una parte de la población estudiada ha sido evaluada una media de 4 meses después de la A. H., otro grupo inmediatamente después, y un tercero y último al comienzo y al final de su A. H. con un test, y al final con un segundo Test diferente.

Los resultados se han valorado mediante procedimientos de estadística inferencial como el Análisis de la Varianza para comparaciones de medias de series no apareadas procedentes de varios subgrupos de la población estudiada. Hemos recurrido al estadígrafo t de Student para comparaciones simples de medias previa comprobación de la hipótesis de igualdad de sus varianzas mediante la prueba de Snédecor o de Cochran. Cuando se ha tenido que recurrir a pruebas no paramétricas de comparaciones de medias muestrales, hemos utilizado el estadígrafo de Mann-Whitney.

En todos los Tests se valora, junto al número de respuestas acertadas, corrección de probabilidad de acierto por azar, el índice de fiabilidad de cada prueba, los índices de dificultad y discriminación de cada pregunta en cada examen, y al margen hemos calculado un nivel aceptable de respuestas antes de la valoración de los resultados, y todo ello mediante los procedimientos habituales.

**C) Estudios de Evaluación Certificativa.** Se ha valorado el cumplimiento de los objetivos registrados en los cuadernos de actividad hospitalaria de 54 alumnos, mediante la construcción de unas escalas de clasificación, que se valoran mediante procedimientos de Estadística Descriptiva.

**D) Análisis de los gastos del proceso docente.** Hemos recurrido a dos procedimientos:

1– Estudio de los rendimientos asistenciales del Servicio durante los meses del Curso Escolar, desde su apertura en el año 1978 hasta la actualidad, con la finalidad de averiguar si existen diferencias con significación estadística de menor rendimiento del Servicio de Cirugía General, desde su dedicación docente, utilizando estadígrafos ya mencionados de comparación de tre variables: número de altas/mes, número de operaciones de cirugía mayor, media y menor/mes, número de enfermos vistos en Consultas Externas/ mes.

2– Análisis de los gastos en material docente, invertido (vestimenta quirúrgica, material para suturas..., usado y desechado (guantes, mascarilla, calzas, hilos de sutura, bisturí,...) y costes de los procesos de limpieza de material quirúrgico. Se valora el coste del personal docente y del material impreso que se le facilitó al estudiante. No se ha valorado determinado material impreso que costeó el estudiante, ni los desplazamientos, ni los costes inmobiliarios.

## RESULTADOS

A) Los resultados de las Encuestas de los estudiantes y pacientes, que reflejan la opinión de estos, en función de unas variables, se ofrecen en las Comunicaciones respectivas en forma de Poster tituladas: Metodología para la Valoración de la Actividad Hospitalaria Pregrada en Cirugía: A) Encuesta a pacientes e idem B) Encuesta a Estudiantes, por lo que se prescinde de su exposición en el presente resumen.

B) El resultado del Análisis de los Tests practicados demuestra que al final de la Actividad Hospitalaria, existe una media de acierto de  $16.56 \pm 0.23$  EE. No existen diferencias estadísticamente significativas ( $p=0.01$ ) entre el grupo examinado inmediatamente después, y el grupo examinado 4 meses después. Existe una diferencia significativa en cuanto al rendimiento en el Test inicial con respecto al final de la A. H. ( $p = 0.01$ ).

No hay diferencia significativa si al final se aplica el mismo tipo de Test que al inicio o uno diferente.

Las medias de los Cursos 1983-84 ( $17.75 \pm 0.26$  EE) y 1984-85 ( $15.82 \pm 0.27$  EE) son estadísticamente significativas ( $p=0.01$ ).

C) Los estudiantes han cumplido el 100% de los objetivos establecidos en las Áreas de Hospitalización y Quirófanos. Todos ellos han entregado ficha resumen de A. H. en forma de Cuaderno de Prácticas. Sus calificaciones certificativas están comprendidas en un intervalo entre el 80.82 y el 100%. La evaluación de cantidad y calidad del trabajo realizado sitúa a los estudiantes en un intervalo comprendido entre el 66.66% y el 100% del que se ha considerado como ideal.

D) 1– No se encuentran diferencias estadísticamente significativas en los parámetros de rendimiento mencionados, cuando se comparan las medias mensua-

les en dos series de número de altas, operaciones y enfermos atendidos en consultas externas.

2- El análisis de los gastos refleja un coste global de 1629410 ptas, lo que supone un coste de 14168 ptas/alumno y 189 ptas/estudiante/hora.

## DISCUSIÓN

De acuerdo con los resultados del estudio, en el apartado de encuestas, nuestro organigrama para la A. H. en Cirugía es necesario y efectivo, para nuestro Servicio, y para los estudiantes, aunque debiera ser modificado puntualmente en la duración de la A. H. Ninguna A. H. sería eliminada por los alumnos. De acuerdo con los resultados de los tests, los estudiantes adquieren unos conocimientos básicos, que anteriormente no habían aprendido. Probablemente no olvidan lo que han aprendido porque durante el curso, realizan A. H. en materias afines, hecho comprobado en estudios de la década de los setenta por Schwartz, G, F Gonella, J, S. y Menduke. H (1974). Con el aprovechamiento de 75 horas A. H. en tres semanas, habitualmente se cumplen los objetivos educativos. El CUADERNO DE A. H. es un método de Evaluación útil, de una A. H. que se ha demostrado efectiva en cuarto curso.

La actividad docente no ha disminuido el rendimiento asistencial. No poseemos datos para comparar los costes del presente estudio, pero pensamos que es necesario reflejar los costes cuando se valora un proceso educativo.

## REFERENCIAS BIBLIOGRÁFICAS

- DEAN, R. E.; JOHNSON T. M. (1977): «The surgical Clerkship in the Community Hospital.» *Journal of Medical Education*, Vol. 52, 59-65.
- DE PALMA, R. G. et alii (1974) «Objetives and methods in undergraduate surgical education». *Surgery*, 75/6, 915-924.
- HAYMAN J. L. (1971). *Investigación y Educación*. Paidós Educador, Barcelona.
- SCHWARTZ, G. F.; GONNELLA, J. S.; MENDUKE H. (1974): «The value of pretesting in the Surgical Curriculum». *Journal of Surgical Research*, 16/2, 131-133.
- VAN DALEN, D. D.; MEYER, W. J. (1981). *Manual de técnica de la investigación educacional*. Paidos Educador, Barcelona.

## Evaluación del rendimiento en el Ciclo Medio de la E.G.B.

por  
*Eduardo García Jiménez*

### CONTEXTO DEL PROBLEMA

Con la publicación en estas fechas del Anteproyecto para la Reforma de la 2.<sup>a</sup> etapa de EGB (1), está en vías de completarse al estructura curricular iniciada en este nivel, cuatro años antes, bajo la denominación de «Programas Renovados». La promulgación de estos programas por el Ministerio de Educación y Ciencia supuso una nueva estructuración de la EGB que afectó, fundamentalmente, a la teleología de la enseñanza y a la selección y secuenciación de los contenidos.

No obstante, el elevado índice de fracaso escolar detectado en la evaluación del ciclo inicial, «tan sólo un 60 por ciento de los alumnos alcanza el 40 por ciento de los aprendizajes que se supone debe proporcionar este primer ciclo de EGB» (2), hace que se esté cuestionando la calidad de la enseñanza consecuencia de la aplicación de los nuevos programas, al tiempo que se inicia una reforma que, presumiblemente, afectará a todos los niveles del sistema educativo.

En general la innovación fue acogida con preocupación por autores y personalidades relevantes vinculadas a la reforma (3) dado que se había establecido un modelo curricular para la EGB sin contar con un programa de evaluación que aportase información sobre el desarrollo de las Nuevas Orientaciones de la EGB, acerca de los errores conceptuales que aún persistían en los alumnos, sobre sus ideas preconcebidas, en definitiva, sobre los esquemas conceptuales de alumnos a los que se les iba a exigir determinados niveles de rendimiento. De un modo particular esta problemática aparecía acentuada en áreas como la de las Ciencias de la Naturaleza (4) que, permanecía sustraída, aún, a cualquier forma de innovación en sus objetivos, contenidos y metodología. Sobre todo, en las primeras etapas de la escolarización, la experiencia directa del alumno con los fenómenos físico-naturales o con el entorno ecológico no había sido fomentada en programas educativos de legislaciones anteriores.

Desde otra perspectiva, la interpretación de la reforma por los profesores estaba relacionada con los cambios que el nuevo currículum obligaría a introducir en el diseño de la instrucción. Sobre todo, cuando estos cambios deberían adaptarse a las necesidades derivadas de una cultura, unas condiciones económicas, sociales y políticas y unas formas de vida diferentes que se oponían a la idea de una clase-patrón, presente en la filosofía de los nuevos programas, en la que los niños cumplen fielmente con los objetivos que se les señala según un determinado modelo de desarrollo que identifica estadios de madurez cognitiva con períodos fijos de escolaridad.

Por todo ello, y antes de que el ciclo superior de la EGB quedase cerrado y se completase así el proyecto curricular para la escuela, parecía necesario: de una parte, llevar a cabo una evaluación diagnóstica que permitiese fijar el nivel de rendimiento alcanzado por los alumnos al finalizar el ciclo medio, para conocer las bases que deberían ser tenidas en cuenta al adoptar las decisiones sobre el nuevo currículum (5); de otra, presentar una propuesta de diseño intruccional en la que sean los profesores los que elaboran su propio material instuccional adaptado al contexto en el que se desarrolla la enseñanza y al ambiente de aprendizaje de su clase.

El trabajo que a continuación presentamos, comprende la evaluación del nivel de rendimiento alcanzado por los alumnos al finalizar el ciclo medio de la EGB en el Área de las Ciencias de la Naturaleza y una propuesta para el diseño intruccional del ciclo superior en este mismo área.

## OBJETIVOS

Los objetivos que nos propusimos cubrir en nuestro trabajo fueron:

- a. Proporcionar información útil a los profesores sobre los errores conceptuales persistentes en sus alumnos, acerca de las actitudes de éstos hacia los contenidos y problemas de la ciencia y sobre un grado de dominio de la metodología científica.
- b. Proporcionar información útil, a los que toman las decisiones curriculares, sobre el grado de dominio alcanzado por los alumnos en los Niveles Básicos de Referencia establecidos para el ciclo medio de la EGB en el Área de las Ciencias de la Naturaleza en función de variables contextuales.
- c. Diseñar un modelo de desarrollo instruccional que permita a los profesores adaptarlo a sus necesidades en el aula.

## HIPÓTESIS DE TRABAJO

Hipótesis Número Uno: La prueba elaborada en nuestra investigación, el «Test Diagnóstico en Ciencias Naturales», es un instrumento fiable y generalizable.

Hipótesis Número Dos: Los resultados en el rendimiento y en las actitudes de los alumnos de la 2.<sup>a</sup> etapa de E.G.B. en el Área de las Ciencias de la Naturaleza, se modifican significativamente en función de la incidencia de determinadas variables contextuales, pero estos cambios no están relacionados con la edad o el sexo de los alumnos.

Hipótesis Número Tres: La administración del «Test Diagnóstico en Ciencias Naturales» facilita la detección de los tipos de errores conceptuales cometidos por los alumnos.

## METODOLOGÍA

Nuestra investigación es descriptiva y se ocupa, por tanto, de presentar los resultados del aprendizaje en ciencias de alumnos de la 2.<sup>a</sup> etapa de la EGB y de la EGB y de descubrir las asociaciones estos resultados con la edad y el sexo de los alumnos y con algunas variables contextuales.

Para ello, hemos procedido a la elaboración de una prueba, el «Test Diagnóstico en Ciencias Naturales», en función de las percepciones de una muestra de 59 profesores de Sevilla y Huelva sobre un banco de ítems ordenados de acuerdo con las categorías y subcategorías de la Taxonomía de Klopfer y con las enseñanzas mínimas establecidas para el ciclo medio de la E.G.B. en el Área de las Ciencias de la Naturaleza (6).

Este instrumento se ha aplicado, posteriormente, a una muestra de niños y niñas de 6.<sup>º</sup> de EGB pertenecientes a cuatro comarcas de la provincia de Huelva, sobre las que se había realizado con anterioridad un análisis socioeconómico y educativo para determinar posibles vías de contraste.

Para determinar la incidencia de estos factores en el rendimiento y en las actitudes de los alumnos, aplicamos el análisis de varianza por rangos de Kruskal-Wallis a las puntuaciones obtenidas en el «Test Diagnóstico en Ciencias Naturales», considerando las 22 unidades muestradas y las cuatro comarcas geográficas en las que están ubicadas. La incidencia de las variables edad y sexo sobre los resultados del aprendizaje es contrastada aplicando la «U» de Mann-Whitney, y se utilizó el análisis de frecuencias para identificar tipos de errores cometidos por los alumnos. Por último, hemos estimado la fiabilidad del test empleando el método de las dos mitades o coeficiente de consistencia interna de Sperman-Brown; su validez fue determinada por el método de análisis de ítems; y, su índice de dificultad se fijó a partir del número de alumnos que contestó correctamente a cada una de las preguntas de la prueba.

## RESULTADOS Y CONCLUSIONES

Los resultados del análisis de la varianza, recogidos en la Tabla N.<sup>º</sup> 1, permiten inferir diferencias estadísticamente significativas entre las siguientes co-

marcas de la provincia de Huelva sujetas a comparación: 1. Huelva capital y La Costa; 2. La Sierra y La Costa; 3. La Sierra y el Área de Huelva; 4. La Sierra y Huelva capital. En estos casos, el valor obtenido al aplicar la fórmula de aproximación

$$Z_a = \frac{N(N+1)}{12} \left( \frac{1}{n_i} + \frac{1}{n_h} \right) \quad \text{para cualquiera de los pares de zonas consideradas es superior a la diferencia de rangos medios } \bar{R}_i - \bar{R}_h \text{ para dichas zonas.}$$

ZONAS CONTRASTADAS	Valor de Za	Valor de $\bar{R}_i - \bar{R}_h /$
HUELVA C.-COSTA	41.296958	55.46497
COSTA-SIERRA	108.97715	191.05518
SIERRA-A. HUELVA	120.27679	176.68848
SIERRA-HUELVA C.	107.78871	135.59021
COSTA-A. HUELVA	67.479370	14.36870
HUELVA C.-A. HUELVA	66.287563	41.09627

TABLA N.º 1. Resultados del contraste establecido entre las cuatro zonas muestreadas de la provincia de Huelva.

En la Tabla N.º 2 se recogen los porcentajes de respuestas incorrectas en los diferentes Temas de trabajo, establecidos para el Ciclo Medio de la EGB en el Área de las Ciencias de la Naturaleza, dadas al «Test Diagnóstico en Ciencias Naturales» por una submuestra de 199 alumnos de la provincia de Huelva.

Como puede observarse en esta tabla, el porcentaje más elevado de respuestas incorrectas corresponde al Tema de trabajo «La Nutrición», con un 75,54% de errores cometidos por los 199 alumnos a los que se les pasó la prueba. En este Tema de trabajo, las conductas más deficitarias detectadas en las respuestas de los alumnos hacen referencia a las subcategorías: «Conocimiento de la terminología científica», «Conocimiento de las tendencias y secuencias» y «Conocimiento de las clasificaciones, categorías y criterios», recogidas en la Taxonomía de Klopfer.

Estos resultados nos acercan a tres conclusiones fundamentales: primera, los alumnos de la 2.ª etapa de E.G.B. no alcanzan los niveles mínimos establecidos para el Ciclo Medio en el Área de las Ciencias de la Naturaleza; segunda, las diferencias socioeconómicas y educativas entre zonas geográficas son un factor que incide en los resultados del aprendizaje de las ciencias tanto a nivel de rendimiento como de actitudes de los alumnos; y tercera, los errores cometidos por los alumnos de la 2.ª etapa de E.G.B. afectan no sólo al ámbito de los conocimientos sino también al de su aplicación a los problemas científicos.

ITEMS del TDCN	TEMAS DE TRABAJO CIENCIAS NATURALES	PORCENTAJE DE ERROR COMETIDO	CONDUCTAS MAS DEFICITARIAS
1,2,3,4, 5,6	1.2. «La nutrición»	75,54	Conocimiento de la terminología, de tendencias y secuencias y de clasificaciones y categorías.
7,8,9, 10,	1.3. «La reproducción»	41,70	Conocimiento de los hechos de la ciencia y de tendencias y secuencias.
11,12	2.2. «Los vegetales»	50,50	Formulación de generalizaciones y conocimiento de los hechos específicos
13,14,15,16	2.3. «Otros elementos y factores del medio: Fuentes de energía	61,05	Adquisición de destrozas en la utilización de equipos de laboratorio
17,18,19	3.1. «Estudio de un ecosistema»	39,36	Conciencia de las relaciones entre la ciencia, la economía y la tecnología.
20,21	3.2. «Aprovechamiento de las fuentes de energía	55,77	Conciencia de las relaciones entre la ciencia... y Aplicación de conocimientos
22,23, 24, 25 26	4. «Técnicas de trabajo»	67,73	Selección de instrumentos apropiados y de pruebas para comprobar la hipótesis.

Tabla N.<sup>o</sup> 2. Resultados del análisis de frecuencias en relación con el porcentaje de error cometido y con las conductas más deficitarias halladas en las respuestas de los alumnos al «Test Diagnóstico en Ciencias Naturales» (TDCN).

#### REFERENCIAS BIBLIOGRÁFICAS

- (1) DIRECCIÓN GENERAL DE EDUCACIÓN BÁSICA, (1984) «Reforma: El Ciclo Superior de la EGB», *Vida Escolar*, N.<sup>o</sup>. 229-230.
- (2) MARAVALL, J. M. (1984): *La reforma de la enseñanza*, Barcelona, Laia, pág. 87.
- (3) UNIVERSIDAD AUTONOMA DE MADRID. (1981) *Los programas renovados de la EGB. Análisis, críticas y alternativas*, Madrid, ICE, 1981; ARRIBAS, C. (1982) *El ciclo medio en la Educación Básica*, Madrid, Santillana; MARTIN, E. y MORENO, A., (1982) «Evaluación y crítica de los nuevos programas de EGB», *Infancia y Aprendizaje*, N.<sup>o</sup> 14.
- (4) OTERO GUTIERREZ, J. (1982) «La fundamentación de los Programas Renovados del Área de las Ciencias de la Naturaleza del Ciclo Medio de la EGB», *Bordón*, N.<sup>o</sup> 242-243, Marzo-Junio, págs. 297-308; SANCHEZ REAL, J. (1982), «Física y Química de los Programas Renovados de la EGB», *Educadores*, N.<sup>o</sup> 116, Enero-Febrero, págs. 31-56.
- (5) TABA, H. (1974), *Elaboración del currículo. Teoría y práctica*, Buenos Aires, Troquel, pág. 305.
- (6) KLOPFER, L.E., «Evaluación del aprendizaje en ciencias», en B.S. BLOOM y otros, (1975) *Evaluación del aprendizaje*, Buenos Aires, Troquel, Tomo III, págs. 93-220.

## El rendimiento académico en la UNED. Un modelo de análisis

por  
*José Luis García Llamas*

El modelo que presentamos a continuación sobre el rendimiento académico en la enseñanza superior a distancia, se basa fundamentalmente en variables de tipo personal de carácter sociológico, pedagógico y didáctico.

Mediante este modelo de análisis intentamos dar respuestas al siguiente interrogante ¿Cómo influyen en el rendimiento académico: la valoración del sistema de enseñanza a distancia, las técnicas y métodos de estudio empleados, la valoración del sistema de evaluación de la UNED y las características personales de alumnos? o dicho de otra forma ¿qué variables inciden en el éxito o fracaso académico de los alumnos de Ciencias de la Educación que estudian en la UNED?

Desde esta perspectiva global, pretendemos diseñar un modelo de análisis en que se recojan los distintos aspectos que pueden incidir en la enseñanza superior a distancia. Al mismo tiempo, intentamos establecer relaciones más específicas entre el éxito o fracaso académico, como variable dependiente y, una serie de variables independientes, con el fin de llegar a conocer en qué medida pueden explicar, condicionar o en otros casos ser condicionadas por el mismo.

Consideramos pues, el rendimiento académico como el producto del proceso de enseñanza-aprendizaje y, más concretamente, el resultado de la aplicación de unas determinadas técnicas o métodos de estudio y trabajo. Gran parte de los investigadores tienden a coincidir en señalar lo impreciso que resulta indentificar el rendimiento académico con las notas; no obstante, también se acepta como su mejor indicador. Así el INCIE (1975) considera que la mejor medida del rendimiento es la puntuación media de las evaluaciones del curso. En la misma línea se manifiestan Carabaña (1979), Pérez Serrano, G. (1981), Rodríguez Espinar, S. (1982) y Molina, S. y García, E. (1984).

Tras una profunda revisión de la literatura relacionada con el tema, podemos afirmar que existen muy pocas investigaciones aplicadas al ámbito de la enseñanza superior a distancia. Entre las más próximas a nuestro objetivo pode-

mos citar la llevada a cabo por el ICE de la Universidad de Salamanca (1981) en la que se destacan la importancia de la orientación del profesor, la actitud de los alumnos, su opinión sobre la enseñanza universitaria y la preferencia casi absoluta por los exámenes parciales frente a los finales. Por su parte Díaz Allúe, T. (1973) subraya las dificultades de los universitarios madrileños en lo referente a técnicas y hábitos de estudio, así como la lectura restringida de obras y autores.

El ICE de la UNED (1981) en una investigación sobre los primeros licenciados analiza la influencia en el rendimiento de los motivos que llevan al alumno a comenzar sus estudios, los logros alcanzados, la asistencia al Centro Asociado, contactos con el profesorado y la utilización de los recursos puestos a su servicio. En esta misma investigación se señalan diferencias de rendimiento en relación a la variable sexo o favor del hombre, también en el nivel de ocupación de los alumnos, en los estudios previos realizados y en la carrera cursada.

Las consideraciones teóricas, indicadas anteriormente, nos han servido de pauta para el diseño de nuestro modelo de análisis, que se circunscribe a los siguientes campos:

- Técnicas de estudio.
- Valoración del sistema de enseñanza a distancia.
- Valoración del sistema de evaluación.
- Características personales.

Estos aspectos se relacionan con el rendimiento académico alcanzado por los alumnos y, medido a través de las calificaciones finales obtenidas en Junio o Septiembre. En el siguiente gráfico presentamos las interacciones e influencias entre las distintas variables contempladas en el estudio.

La muestra utilizada en este trabajo se seleccionó, inicialmente, mediante el muestreo probabilístico sistemático. Está compuesta por 805 alumnos de Ciencias de la Educación distribuidos por todo el territorio peninsular e insular, incluso alguno de los alumnos residía en el extranjero. Como instrumento de recogida de datos, se ha utilizado un cuestionario de opinión sobre aspectos didácticos y pedagógicos de la UNED, elaborado específicamente por nosotros para este estudio.

Para facilitar el tratamiento estadístico de los datos hemos utilizado el paquete de programas del S.P.S.S., como técnicas estadísticas más destacables citaremos: *Ji cuadrado*, análisis de varianza, análisis discriminante.

## PRINCIPALES RESULTADOS

En primer lugar, cabe resaltar que son muy pocos los alumnos que se presentan a todas las asignaturas de las que se han matriculado, especialmente cuando se trata de alumnos que cursan su primer año en la UNED, primer curso o curso de Adaptación para profesores EGB.

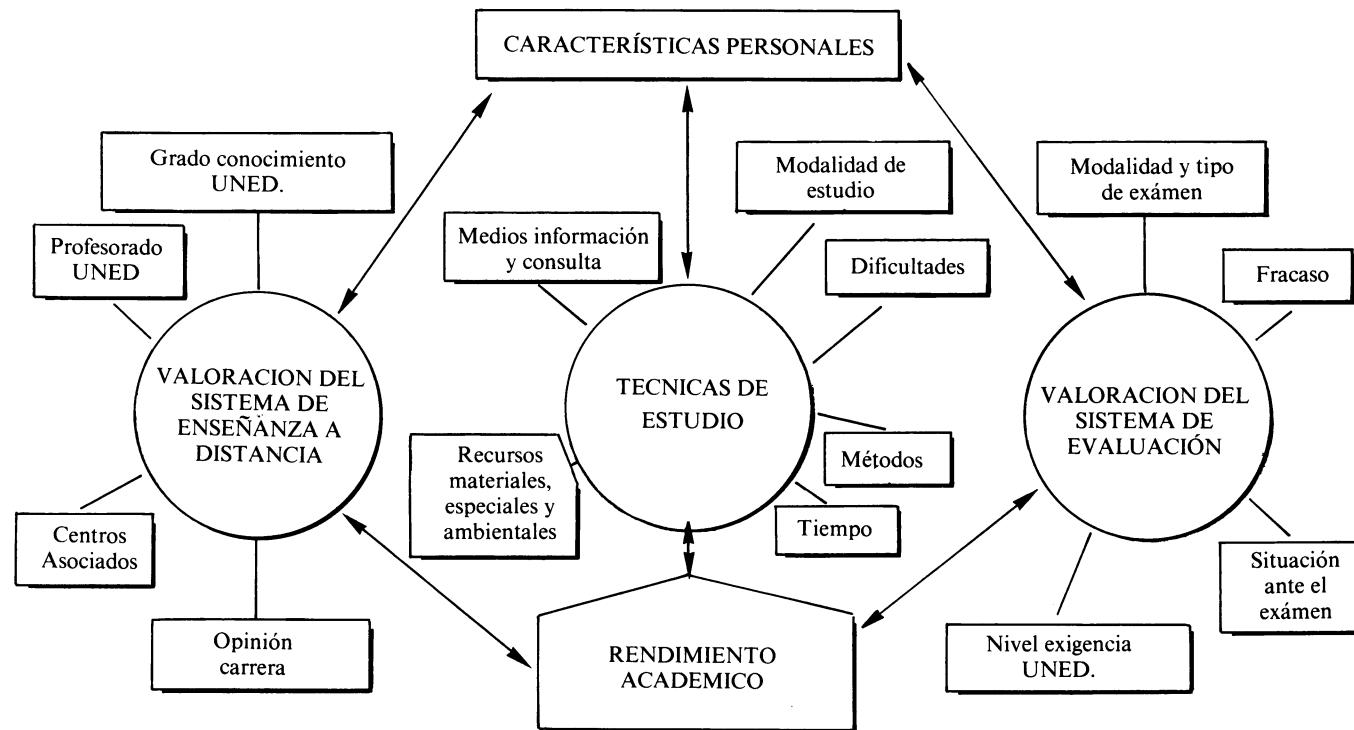


Gráfico n. 5.1: Variables consideradas en el estudio.

Consideradas globalmente las calificaciones obtenidas por los alumnos, teniendo en cuenta los que se han presentado a examen en alguna de las materias –un 65,6% lo hace a tres o menos asignaturas–, observamos que un 24% no se ha presentado a examen, un 6,7% ha obtenido la calificación de suspenso, un 47,8% logra el aprobado, un 18,3% obtiene el notable y tan sólo un 2,4% alcanza al- canza el sobresaliente. En general, se pone de manifiesto que las calificaciones de la UNED son bajas.

Por su parte, las materias que suele superar un mayor número de alumnos son las de tipo didáctico, metodológico y teórico, mientras que las asignaturas de tipo histórico y sobre todo experimental encierran, por lo general, para el alumno una mayor dificultad y, por tanto, son superadas por un menor número de alumnos.

Tras estas consideraciones previas, nos interesa resaltar, especialmente, algunos de los rasgos distintivos que identifican a los alumnos de Ciencias de la Educación que alcanza el éxito en sus estudios, de esta forma podemos diseñar el perfil de este tipo de alumno en la enseñanza superior a distancia, las notas definitorias de este grupo son las siguientes:

- Suelen tener una opinión favorable sobre la UNED y su sistema de enseñanza.
- Mantienen contactos frecuentes con el profesorado, este contacto se realiza mediante la correspondencia y el teléfono con el profesor de la Sedé Central y de forma personal con el profesor-tutor del Centro Asociado.
- Cursan estudios universitarios para alcanzar el prestigio social que proporciona la obtención del correspondiente título, así como para aplicar los conocimientos a su profesión y enriquecer su propio autoconcepto.
- Tienen una opinión positiva sobre la carrera que cursan y los contenidos que en ella se imparten.

Emplean en su estudio más de tres medios de información y, además, de las Unidades Didácticas y los libros obligatorios, utilizan los libros recomendados, consultan los diccionarios especializados, escuchan las emisiones de radio y las cassettes, a la vez que leen tres o más revistas especializadas en temas educativos.

- Al efectuar su estudio independiente realizan con asiduidad las siguientes actividades: relacionar unas materias con otras, leer todo el tema para tener una visión global del mismo, seleccionar los conceptos básicos del tema y elaborar fichas con los conceptos básicos.
- Suelen dedicar al estudio de las asignaturas doce o más horas semanales, además alcanzan un alto grado de concentración en los estudios.
- Sienten una gran inquietud por los temas educativos, por ello, suelen adquirir un mayor número de libros que los estrictamente obligatorios. En general, disponen de una habitación para el estudio en la que están solos. Aunque el ambiente de la casa no favorece demasiado su estudio, sus familiares les animan a estudiar y, además, suelen hacerlo a las mismas horas.

- Manifiestan tener muy pocas dificultades en técnicas y métodos de estudio y trabajo personal.
- Prefieren las pruebas presenciales basadas en los temas de ensayo y los comentarios de texto, frente a la prueba objetiva y las preguntas breves.
- Opinan que las causas del fracaso en los exámenes se deben principalmente, en el caso de la UNED, a la amplitud de los programas exigidos, al cansancio que producen los exámenes y a la falta de rigor y sistematización en los estudios.
- Suelen ser muy autocríticos en la valoración de su rendimiento personal.
- Predominan los varones sobre las mujeres y las edades maduras 30-38 años y jóvenes, menos de 22 años, frente a las edades intermedias 23-29 años.

En síntesis, como resultado de aplicar la técnica del análisis discriminante a las variables del estudio, hemos podido comprobar que los rasgos con mayor poder discriminatorio en los alumnos con éxito, son los que se indican a continuación, siguiendo el orden puesto de relieve en el análisis:

- autovaloración personal
- tiempo semanal dedicado al estudio
- frecuencia de las consultas al profesor-tutor
- capacidad para relacionar unas materias con otras
- utilización y disponibilidad de un buen número de recursos materiales
- medios empleados en la preparación de las asignaturas y
- opinión sobre la carrera de Ciencias de la Educación.

Por tanto, estos rasgos nos permiten identificar el perfil del alumno con éxito en la Universidad en el sistema de enseñanza superior a distancia.

#### REFERENCIAS BIBLIOGRÁFICAS

- BENZECRI, J.P. (1984): *L'analyse des données*, Ed. Dunod, Paris.
- CARABAÑA, J. (1979): «Origen social, inteligencia y rendimiento académico al final de la EGB». En *Temas de investigación educativa I*, Ed. MEC, Madrid.
- CIRIGLIANO, G. (1983): *La educación abierta*, Ed. El Ateneo, Buenos Aires.
- DANIEL, J.S. (1984): «The future of distance teaching Universities in a worldwide perspective». En *Evaluación del rendimiento en la enseñanza superior a distancia*. Ed. UNED, Madrid.
- DIAZ ALLUE, T. (1973): *Problemática del universitario madrileño*, Ed. CSIC, Madrid.

- GARCÍA LLAMAS, J.L. y PÉREZ SERRANO, G. (1983): *Los Centros Asociados. El Centro de Madrid*, Ed. UNED, Madrid.
- HERRERO CASTRO, J.: INFESTAS GIL, D. (1980): *El rendimiento académico en la Universidad*, ICE de la Universidad de Salamanca.
- HOLMBERG, B. (1985): *Educación a distancia: situación y perspectivas*, Ed. Kapelusz, Buenos Aires, Argentina.
- KLECKA, W.R. (1980): *Discriminant analysis*, Sage University Paper 19, Sage Publications, Beverly Hills.
- MARÍN IBÁÑEZ, R. (1984): «El sistema pedagógico de la UNED y su rendimiento». En *Evaluación del rendimiento de la enseñanza superior a distancia*, Ed. UNED, Madrid.
- MOLINA, S. y GARCÍA, E. (1984): *El éxito y el fracaso escolar en la EGB*, Ed. Laia, Barcelona.
- PÉREZ SERRANO, G. (1981): *Origen social y rendimiento escolar*, Ed. CIS, Madrid.
- RODRÍGUEZ ESPINAR, S. (1982): «Un modelo de predicción del rendimiento académico en la segunda etapa de EGB». En *Modelos de investigación educativa*. ICE de la Universidad de Barcelona.
- UNED. (1981): *Los primeros licenciados de la UNED*, ICE de la UNED, Madrid.
- VARIOS. (1976): *Determinantes del rendimiento académico*, INCIE, Publicaciones del MEC, Madrid.
- (1984) *Evaluación del rendimiento de la enseñanza superior a distancia*. Ponencias presentadas al Congreso Internacional de Universidades Abiertas y a Distancia celebrado en Madrid, Octubre de 1983.

## La eficacia docente de la UNED: aproximación a un modelo

por  
*Carmen Jiménez Fernández*

### INTRODUCCIÓN

Ya no cabe plantearse la posibilidad (Marín Ibáñez, R. 1980, Sarramona, J. 1975, 1980) y relevancia (Sarramona, J. 1981, Jiménez Fernández, C. 1985) de la educación a distancia. Sin embargo, hay que admitir que en el caso de la Universidad Nacional de Educación a Distancia (UNED) se trata todavía de una pedagogía experimentada más que experimental (García Hoz, V. Pérez Juste, R. 1984) a pesar de sus ochenta mil alumnos. El análisis objetivo de su eficacia docente debe continuar y mejorarse.

Dicho análisis implica varios problemas. Si como indica Ferguson, J. (1975) «la eficacia de una universidad se juzga por sus graduados», ¿cómo definirla y medirla?. ¿Se puede tomar el rendimiento académico del alumno como indicador global de eficacia? (Jimeno, J. 1981). Dado el índice de abandono (el 60% antes de comenzar hasta el curso 1981-82) y las características socioprofesionales de los que continúan (ICE, 1979-80/81-82; Juliá, S. 1984) y culminan los estudios (ICE, 1981), ¿cabe pensar que la UNED juega un escaso papel en la dinamización del contexto social y personal del alumno? ¿Cuál es el profesor ideal en la universidad (Aparicio, J. San Martín, R. Tejedor, J. 1982) a distancia? ¿Y el tutor ideal?. Omitimos más interrogantes. Aunque existen investigaciones sobre ella (Jiménez Fenández, C. y Pérez Juste, R. 1985), aportan datos más para reconducir la investigación que para pautar la actividad académica.

### NECESIDAD DE UNA TEORÍA

Necesitamos una teoría de la educación superior a distancia que ofrezca esquemas conceptuales y estructuras teóricas que oriente las hipótesis, sugiera relaciones y permita probarlas en la realidad (Kerlinger, F.N. 1975). La investigación realizada sin referencia a un modelo teórico es cuestionada (Seoane, 1979).

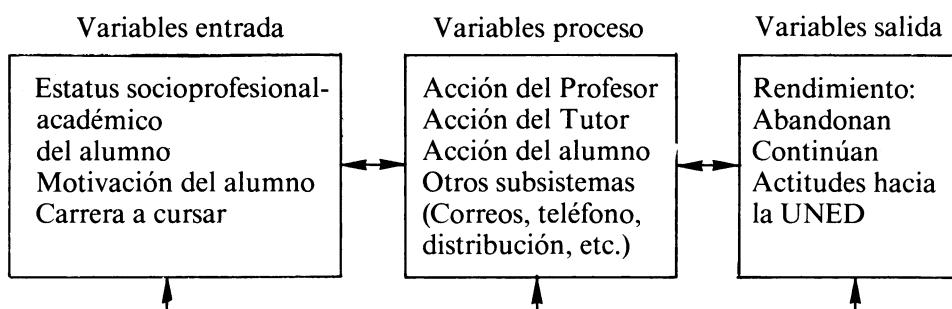
Una aproximación a esta línea es la utilización de modelos que pueden servir de filtro entre teoría y realidad, pensamiento y acción (Bunge, N. 1975).

Escudero Muñoz (1980) expone una buena síntesis de modelos de investigación, y otros autores (Stufflebeam y otros, 1971; Jimeno, J. 1985; Río, P. y otros, 1985) defienden asimismo su necesidad. El propuesto por Escudero Muñoz, J. N. (1982) es válido para nuestros fines porque: a) Incluye la casi totalidad del conjunto de variables de otros modelos (*inputs-outputs*; proceso-producto; mediacional, ecológico), b) Admite variables específicas y sugiere relaciones propias del contexto-proceso donde se trabaja y c) Es fácil ver cuáles han sido analizadas, con qué metodología y que efecto han tenido sobre la investigación y la acción.

Se parte de una realidad, la UNED y de un modelo educativo, el implementado por ella.

## FORMALIZACIÓN DEL MODELO

Considera tres conjuntos de variables:



Resaltar en primer lugar la complejidad y multiplicidad de interacciones que sugiere tanto en el interior de cada conjunto como entre ellos. En segundo lugar creemos que las relaciones y peso de las variables puede ser diferente a las habidas en la universidad convencional. Hay variables de entrada exógenas a la institución (estudios con los que se accede, continuidad o no entre estudios previos y acceso a la Universidad, etc.) y endógenas (existencia o no de Centro Asociado y/o tutor) que pueden ser covariantes importantes de la decisión abandono/continuidad (Jiménez Fernández y otros 1983) y del rendimiento académico, sobre todo en los primeros cursos. En las variables proceso, características de los profesores consideradas importantes tradicionalmente, sólo ocasionalmente puede influir en el seno de la UNED. Así las relaciones con el factor verbal (claridad y fluidez expositivas, por ejemplo) y direccional del aprendizaje (entusiasmo en clase, refuerzos verbales y gestuales, etc.). En cambio, cobran relevancia las relacionadas con su *capacidad para estructurar y trasmitir por escrito* (material

impreso, correspondencia con alumnos) o por *medios tecnológicos* (radio, cassette, video) *los mensajes*. En el caso del alumno, el *estudio independiente* eficaz es capital. Son muchos los que no asisten a las tutorías y aunque asistan, el estudio sostenido en solitario es clave para el aprendizaje a distancia. Entre profesor y alumno media el tutor, variable compleja y polémica cuya eficacia (Jiménez Fernández, C, 1983) y funciones están por aclarar.

En cuanto al producto proponemos dos variables. La primera, el rendimiento académico, es controvertida como medida adecuada de la eficacia de un centro (Jimeno, J. 1981, 1985). Sin renunciar a su mejora, hay que considerarla aunque en el contexto que nos ocupa es una medida menos «consistente» debido a lo errático de su aparición (abandono y número de asignaturas a las que se presentan), a su fiabilidad y validez en consecuencia, si se toma como criterio convencional exclusivo y a la escasa variabilidad que arroja. Conviene analizarla dentro de subgrupos y utilizarlo como variable independiente en estudios ex-post-facto. En cuanto a las actitudes no se incluyen sólo por aproximarnos al ámbito efectivo sino también porque los motivos para estudiar en la UNED parecen ofrecer una muy amplia variabilidad (Juliá, S. 1984) que puede estar relacionada con el comportamiento diferencial del rendimiento.

Ante lo anterior cabe preguntarse si la educación a distancia no estará desarrollando un nuevo estilo cognitivo (Aguado Arrese, A. 1980) en alumnos y profesores o más simplemente un nuevo estilo de enseñanza-aprendizaje cuyas condiciones y efectos hay que evaluar. Razones de espacio impiden otras consideraciones metodológicas. Sirva como broche el que los estudios descriptivos, correlacionales y experimentales tienen amplia cabida en este modelo (Alvira, F. y otros, 1979).

## INVESTIGACIONES REALIZADAS

Los trabajos que reseñamos brevemente, y otros más teóricos que podrían incluirse, explican en gran medida la mayor complejidad de los que esperamos dar a conocer en breve.

- 1) **Autor y fecha:** Otero Vázquez, J. S. (1981) (Directora: Jiménez Fernández, C.)

**Título:** *El Centro Regional de la UNED de Pontevedra visto por sus alumnos.*

**Objetivo:** Recabar de los alumnos información directa sobre el funcionamiento académico-cultural del Centro.

**Datos:** Respuestas a un cuestionario elaborado para tal fin.

**Muestra:** 241 alumnos (65 por cien varones).

**Diseño:** Descriptivo de grupo único.

**Análisis:** Descripción de las respuestas en tanto por ciento.

**Resultados:** Entre otros consideran positiva la labor cultural del Centro, los seminarios y conferencia y la figura y acción del Tutor. Prevalecen

las opiniones negativas respecto al trabajo en equipo, a la orientación que reciben antes de matricularse, a la falta de retroalimentación en las pruebas de evaluación a distancia y a los fondos de la biblioteca.

- 2) **Autor y fecha:** Jiménez Fernández, C. (Curso 1981/82).

**Título:** *Orientaciones sobre Pedagogía Diferencial.*

**Objetivo:** Tener información directa y pertinente sobre los puntos fundamentales de la materia que impartía y ver que medios de consulta con el profesor eran más utilizados por tutores y alumnos.

**Datos:** Encuesta anónima enviada a los alumnos presentados al examen de junio. Control de la frecuencia y medios empleados por alumnos y tutores para comunicarse con el profesor de la Sede Central.

**Muestra:** Cien encuestas válidas que coinciden con la casi totalidad de los alumnos presentados a examen.

**Diseño:** Descripción de grupo único.

**Resultados:** Prevalecen las opiniones positivas sobre las Unidades Didácticas, orientaciones dadas, tipo de exámenes y calificación. El 43% recibió del profesor respuesta eficaz y el 8% no. Los medios de comunicación empleados por los alumnos son el teléfono, la carta y la entrevista en orden decreciente y en el caso de los tutores el teléfono, la entrevista y la comunicación postal.

- 3) **Autores y fecha:** (Jiménez Fernández, C.; González García, M.J.; y Pradera San Segundo, S.) (Curso 1982/83).

**Título:** *Aproximación a la escala valorativa de los alumnos que abandonan y continúan en la UNED*

**Objetivo:** Ver si los alumnos que culminan la carrera y los que la abandonan en primero y/o segundo curso, difieren en su escala de valores.

**Datos:** Respuestas dadas al Test de Relación Valorativa.

**Muestra:** 21 alumnos que habían abandonado los estudios y 21 que habían coronado su carrera, localizados a través de los archivos de dos Centros Asociados.

**Diseño:** Dos grupos independientes con postest

**Ánalisis:** Prueba U para muestras grandes y pequeñas. Se analizaron las respuestas de ambos grupos y por sexo atendiendo a la intensidad y al sentido de la reacción valorativa. También se analizaron las respuestas extremas –muy simpático/muy antipático–por si aparecía alguna tendencia significativa.

**Resultados:** De los veinte contrastes realizados con los dos grupos, sólo cuatro arrojaron diferencias significativas. Los contrastes de sólo las respuestas extremas arrojan diferencias significativas en los códigos dinero, actividad y arte en el extremo muy simpático y no aparecen en el extremo muy antipático. En cambio el sexo arroja diferencias significativas tanto en intensidad cuanto en sentido.

4) **Autor y fecha:** Jiménez Fernández, C. (1983).

**Título:** *La heteroevaluación en la UNED y la eficacia de dos medios de formación.*

**Objetivo:** 1) Obtener información contrastada para estimar el peso a otorgar a los informes de los tutores y evaluar la posible eficacia de los seminarios para tutores y de las convivencias. 2) Estudiar el efecto que produce en los alumnos el feedback sistemático sobre los exámenes.

**Datos:** Evaluación global otorgada por profesor, tutores y alumnos a un trabajo práctico obligatorio. Evaluación de aspectos particulares del mismo por parte del profesor y de los tutores. Evaluación de profesor y alumnos a las dos pruebas presenciales, dando éstos la primera autoevaluación después de conocer aciertos y fallos del examen y en el momento de realizarlo en el segundo caso.

**Muestra:** 41 protocolos cumplimientos por los autores y el profesor y 33 cumplimientos por los alumnos, todos pertenecientes a la asignatura de Pedagogía Experimental II.

**Diseño:** «Cuasiexperimental» con dos o tres grupos independientes según el tipo de hipótesis.

**Análisis:** Análisis de varianza y comparaciones a posteriori en el caso del diseño de tres grupos.

**Resultados:** Los autores califican en todos los casos más alto que el profesor y que se autocalifican los alumnos. Estas diferencias son siempre significativas entre profesor y tutor, y sólo en un caso de siete, entre tutor y alumnos. Profesor y alumnos califican de modo similar cuando éstos se autoevalúan después de conocer aciertos y fallos del examen realizado, pero califican de modo significativamente distinto, cuando el alumno se autoevalúa tras realizar el examen. La asistencia al seminario para tutores no contribuye a diferenciar las calificaciones que otorgan a sus alumnos. En cambio la convivencia aparece como un medio eficaz para aproximar los criterios de evaluación de tutores y profesores y diferencia entre tutores asistentes y no asistentes.

#### REFERENCIAS BIBLIOGRÁFICAS

AGUADO ARRESE, A. (1980): «Perspectivas educacionales del pensamiento tecnológico». *Revista de educación* (263), MEC, Madrid.

ALVIRA, F.y otros (1979), *Los dos métodos de las Ciencias Sociales* CIC, Madrid.

APARICIO, J., SANMARTIN, R. y TEJEDOR, J. (1982): «La enseñanza universitaria vista por los alumnos: Un estudio para la evaluación de los profesores en la enseñanza superior» de capacitación docente OEI, Madrid.

- BUNGE, N. (1975): *Teoría y realidad*, Ariel, Barcelona, p. 34.
- ESCUDERO MUÑOZ, J.N. (1980). «La eficacia docente: estudios correlacionales y experimentales» en: *La investigación pedagógica y la formación del profesorado*, CSIC, Madrid, pp. 207-235.
- (1982). «El fracaso escolar. Hacia un modelo de análisis» en: *Modelos de investigación educativa*. ICE, Barcelona, pp. 17-75.
- FERGUSON, J. (1975). *The Open University from within*, University of London Press Ltd., London, p. 149.
- GARCÍA HOZ, V. y PÉREZ JUSTE, R. (1984), *La investigación del profesor en el aula*, Escuela Española, Madrid, p. 19.
- ICE, *La UNED y sus alumnos*. Curso 1979/80; 80/81; 81/82, UNED, Madrid.
- JIMÉNEZ FERNÁNDEZ, C., CONZÁLEZ GARCÍA, J. y PRADA SAN SEGUNDO, S. (1983). «Evaluación de algunas características de los alumnos y de la metodología de la UNED», en: *I congreso de Tecnología Educativa*, SEP, Madrid.
- JIMÉNEZ FERNÁNDEZ, C. (1983) «Evaluación del rendimiento en la UNED». *Revista española de pedagogía* (1982), CSIC, Madrid. pp. 591-606.
- (1985) «La orientación en la Universidad a Distancia» en: *Seminario de Pedagogía Universitaria*. Real academia de Ciencias Morales y Políticas, Madrid, En prensa.
- JIMÉNEZ FERNÁNDEZ, C. y PÉREZ JUSTE, R. (1985), «Evaluación de algunos aspectos curriculares de la UNED», *Revista de investigación educativa*, ICE, Universidad Barcelona, (En prensa).
- JIMENO SACRISTÁN, J. (1981). «La integración de la teoría del aprendizaje en la teoría y práctica de la enseñanza» en: PÉREZ GÓMEZ, A. y ALMARAZ, J., *Lecturas y apredizaje y enseñanza*, Zero-Zyx, Madrid, pp. 465-499.
- (1985) *Teoría de la enseñanza y desarrollo del currículum* Anaya, Madrid.
- JULIA, S. (1984). «Demandas de estudios y resultados académicos en la UNED» en: *Evaluación del rendimiento de la enseñanza superior a distancia*, UNED, Madrid, pp. 453-474.
- KERLINGER, F. N. (1975), *Investigación del comportamiento*, Interamericana, Méjico, p. 3-4.
- MARÍN IBÁÑEZ, R. (1980), «Beneficio y eficacia de la educación a distancia». *Revista de educación* (263) (MEC, Madrid).
- RÍO, P. y otros (1985). «Ecología y educación», *Rev. Infancia y aprendizaje*, 29, Monográfico.
- SARRAMONA, J. (1975) *Tecnología de la enseñanza a distancia*, CEAC, Barcelona.
- (1980) «Problemas y posibilidades de la enseñanza a distancia». *Medios audiovisuales*. (91). Madrid.
- (1981) «Teoría y práctica de la educación a distancia» en: *Seminario Iberoamericano sobre Diseño y Evaluación de programas de Educación Superior a Distancia*, OAE-C, PEIP, Santiago de Chile.

# La investigación sobre el rendimiento desde una perspectiva sociológica

por

*A. López Blasco/J. Marqués Ruiz/A. Martínez Sánchez*

## INTRODUCCIÓN

Para iniciar nuestras consideraciones metodológicas sobre el rendimiento escolar desde una perspectiva sociológica cualitativa, interesa perfilar el concepto de rendimiento que constituye el objeto sobre el que se centra nuestra propuesta. De entre las acepciones dadas por el Diccionario de la Real Academia Española al término «rendimiento», la más cercana al sentido de nuestro estudio es la que lo define como «producto o utilidad que rinde o da una cosa».

En general se habla de rendimiento aludiendo al trabajo realizado en una unidad de tiempo y en esta línea de cociente de rendimiento ( $r = \text{trabajo/tiempo}$ ).

En sentido amplio, rendimiento es el valor producido al aplicar determinada cantidad de energía.

En psicología se habla de rendimiento en referencia a las capacidades del hombre o de cualquier organismo puesto en acción.

En pedagogía el concepto de rendimiento hace referencia a los logros en una determinada tarea previamente establecida. En este sentido subyace una tendencia, explícita o no, hacia la objetividad tanto en el empleo de instrumentos de medida como en la expresión de resultados. La valoración del rendimiento de los individuos se realiza con referencia al criterio y/o a la norma.

Restringirse en el ámbito educativo a esta acepción equivale a aceptar y a entender la enseñanza como un mercado ajustado a las características del capitalismo de producción y consumo de masas, regido por el valor del cambio (Elejaldeitia, C. 1982).

Esta concepción supondría considerar que tanto la oferta como la demanda educativas responden a un único proyecto, el de la producción. Bajo esta rúbrica se encuentra, a nuestro entender, gran parte de la investigación sobre el rendimiento situada en la dimensión del eficientismo.

Sin embargo, entendemos que el único rendimiento significativo es el de la propia vida, es decir, el que se manifiesta a través del propio comportamiento. Y éste se produce sobreañadiendo al «rendimiento escolar» en su acepción reduccionista, las aportaciones de los sujetos en cuanto espectadores críticos del producto de la escuela. Es una aportación que conduce a la necesaria redefinición del «proyecto de rendimiento» que la sociedad establecida, el sistema escolar, los padres y los maestros, proyectan como real oferta/demanda de la educación.

Desde esta perspectiva cobran especial interés las propuestas de investigación cualitativa que tratan de presentar una «reconstrucción de la realidad» teniendo en cuenta el punto de vista del actor, es decir, las que destacan el interés por la subjetividad.

## LA SUBJETIVIDAD COMO OBJETIVO DE INVESTIGACIÓN.

Los principios sociológicos del círculo de Mead, G.H. (1934), del interaccionismo simbólico y de la etnometodología subrayan la importancia del medio social para el actuar cotidiano constitutivo de sentido en interacción con un sujeto que se refiere de modo reflexivo a lo social y que cree él mismo una praxis social.

Los principios estructuristas y psicoanalíticos, en cambio, consideran al sujeto más bien como una estación de tránsito de fenómenos transsubjetivos, por ejemplo, del inconsciente, de la dinámica familiar, de las reglas de interacción. Oevermann, U. (1972) afirma que es característico de los individuos el que se interpretan mal a sí mismos y a sus motivos de acción y no ven las fuerzas que les condicionan.

Para Wahl, K. (1983) es pensable la existencia de grados de subjetividad: subjetividad poco desarrollada, socialmente oprimida o encubierta. Según sus gradaciones en los casos individuales concretos la subjetividad puede abarcar sufrimiento, ideologización, impotencia y resistencia, autoconciencia y poder en acción.

Una teoría del sujeto es siempre también una teoría de su socialización funcional, al mismo tiempo que de su dominio de la vida a la búsqueda de su intensidad.

En el estudio de la problemática de la subjetividad formada en interacción social, aparece la dimensión ético-política. Así se presentan problemas de poder y dominio, respecto de la diferencia, tolerancia con los mundos subjetivos de otros etc.

Aparece este problema cuando, por ejemplo, el mundo subjetivo de los investigados se describe según los criterios culturales del investigador o de aquellos para quienes éste trabaja.

Para las Ciencias Sociales esta cuestión se plantea sobre todo cuando han de entrar al servicio de intereses filantrópicos de promoción social, sociopolíticos o

de ayuda al desarrollo. Se espera de la investigación que estudie las interpretaciones subjetivas de los miembros de los grupos en estudio para obtener de este modo información en orden el planteamiento y desarrollo de «medidas adecuadas».

Por otra parte, ante el desconocimiento de los otros, fácilmente se pueden dar problemas de idealización y de proyección de las propias opiniones.

Pocas veces se intenta comprender la subjetividad de los otros desde ellos mismos. Así es explicable la posición de Peter L. Berger, que con respecto a los problemas del desarrollo del Tercer mundo aboga por el reconocimiento de las correspondientes definiciones culturales y subjetivas de los interesados que para su caso son más adecuadas que las de los políticos extranjeros del mundo desarrollado (Wahl, K. 1983).

La investigación social orientada hacia la subjetividad ha de reflejar con exactitud las líneas de demarcación del poder de definición –hasta qué punto los mundos subjetivos son definidos por elementos externos– y las del poder de intervención –en qué medida los hechos externos intervienen en los mundos subjetivos-. De lo contrario la investigación social se convertiría en cómplice de controles y dominio, sutiles o abiertos, sobre minorías que se desvían de la norma o sobre mayorías subprivilegiadas.

Esta orientación científica connota generalmente un interés por plasmar los resultados en propuestas concretas para transformar la realidad estudiada. Se trata de intervenir, con los datos obtenidos en la discusión pública y política, informando del hecho de que la subjetividad, las opiniones, las convicciones, los valores y las normas de las personas no solo pueden ser considerados como algo «espontáneo» sino que también son «producidos» por las condiciones de vida, los medios de comunicación, la manera con que los otros ven a estas personas y también por los mismos métodos de la investigación.

Creemos que los conocimientos cualitativos, más que las visiones cuantitativas, aparte de ofrecer una mejor posibilidad de conocer determinados problemas pueden ayudar al desarrollo de medidas adecuadas para mejorar la situación de los afectados por el tema.

## **APLICACIÓN DE LA METODOLOGÍA DEL INTERACCIONISMO SIMBÓLICO A UN EJEMPLO CONCRETO SOBRE RENDIMIENTO ESCOLAR INSATISFACTORIO.**

En el marco de la concepción fenomenológica, el Interaccionismo Simbólico desarrolla un esquema analítico de la sociedad y del comportamiento individual humano basándose en la interpretación del significado que las cosas y los hechos tienen para las personas. El universo de símbolos interiorizados por las personas, para los seguidores de esta corriente, surgen de la interacción llevada a cabo son los otros y es así como se va formando el «self»: entre los «estímulos» exteriores y las propias reacciones (Blumer, H. 1969).

Dentro de este paradigma interpretativo se propone un camino a seguir en dos fases: la exploración y la inspección. A través de la exploración se entra en contacto directo con una parte de la vida social de manera flexible, no fijada de antemano y en la que el investigador desde un postura de «observación participante» adquiere un conocimiento de miembro que le proporciona una «definición de la situación» a través de las percepciones e interpretaciones de la realidad en tanto estas se relacionan con los comportamientos de los sujetos. La exploración no va unida a un determinado tipo de técnicas. Se puede echar mano de la observación directa, de la entrevista en profundidad, participación no directiva en conversaciones, diarios personales, etc., cuyo uso no está sujeto a un procedimiento formal previamente establecido.

La fase de la «inspección» intenta un examen intensivo del contenido empírico de cierto número de elementos analíticos ya obtenidos, así como el estudio de la naturaleza de las relaciones entre tales elementos.

En el trabajo realizado sobre el «Fracaso Escolar» (López A., Marquéz J. y Martínez A., 1985) la exploración empírica se ha basado en entrevistas en profundidad con un guión ligeramente estructurado que ofrece la oportunidad al entrevistado de expresarse libremente ante un extraño interesado sobre su forma de entenderse a sí mismo y a la realidad que le rodea. Se proponen dos grandes núcleos de contenido en torno al rendimiento: la consideración general de la escuela en sus aspectos positivos y negativos y las percepciones de los alumnos sobre el hecho del fracaso escolar y su visión sobre los motivos que producen la falta de rendimiento académico y los efectos generados en ellos.

En la etapa analítica, después de la transcripción y lectura de las entrevistas realizadas, hemos tratado de «sumergirnos» en la realidad de los sujetos tal como es descrita y vivida por ellos mismos, tratando de obtener una visión panorámica de cada una de las entrevistas.

Posteriormente se han delimitado las unidades temáticas, identificando las áreas significativas de la estructura del relato en torno a la escuela y al fracaso en ella.

Otro punto ha sido la determinación del significado central dominante a través de las expresiones de los entrevistados buscando la perspectiva en la que se sitúan los datos consciente y/o inconscientemente emitidos. Se realiza aquí un proceso hermenéutico por el que se trata de encontrar los significados que no aparecen en una primera aproximación.

La fase siguiente consiste en la traducción del fenómeno a un lenguaje científico elaborado lo que supone una reflexión sobre los temas centrales para conseguir la traducción de estos núcleos temáticos a un lenguaje universalizable.

A continuación se pasa a una globalización de cada uno de los temas en una estructura general. Se trata de descubrir las estructuras básicas que subyacen al fenómeno en consideración.

Finalmente se intenta lograr una integración de las estructuras parciales en un esquema general que nos permita la determinación de la fisonomía del grupo estudiado en relación a su rendimiento académico. Aquí se consideran las cir-

cunstancias –quienes, qué, dónde, cuándo, cómo, a que nivel y por qué– en las que se sitúa su comportamiento referido al fenómeno social del fracaso escolar.

Teniendo en cuenta que nos encontramos en el marco de un «paradigma interpretativo», que no pretende la medida de variables aisladas, sino la interpretación de los fenómenos a partir de los símbolos lingüísticos, la discusión de los resultados tiene como finalidad comparar, contraponer, las opiniones de los diversos investigadores para llegar a completar y enriquecer el conocimiento del fenómeno estudiado, es decir, una reconstrucción de la realidad interesada en el servicio de las personas afectadas.

#### REFERENCIAS BIBLIOGRÁFICAS

- BLUMER, H. (1969): *Symbolic Interactionism. Perspective and Methods*. Prentice-Hall Inc. Englewood Cliffs. N.J. USA. traducción castellana Ed. Hora, Barcelona 1982.
- ELEJABEITIA, C. (1982): «Fracaso Escolar». *Cuadernos de Pedagogía*, n.º 88, abril. pp. 59-62.
- LÓPEZ A., MARQUES J., MARTÍNEZ A. (1985): *Fracaso Escolar. Percepciones y vivencias de los alumnos*. Institución Alfonso, el Magnánimo. Diputación de Valencia.
- MEAD, G.H. (1934): *Mind, Self and Society*. The University of Chicago Pres. Traducción castellana en Paidós, Buenos Aires, 1967.
- OEVERTMANN, U. (1972): *Überlegungen zum Zusammenhang von Sprache, kognitiver Entwicklung und sozialer Herkunft*. En: Holzer, H./Steinbacher, K., (ed.), *Sprache und Gesellschaft*. Hoffman et Campe, Hamburg, 1972.
- WAHL, K. (1983): *Renaissance der Subjektivität: Sozialforschung und Reformpolitik* (documento policopiado).

## El cine didáctico y los estilos cognitivos Un modelo de análisis interactivo.

por  
*Juan de Pablos*

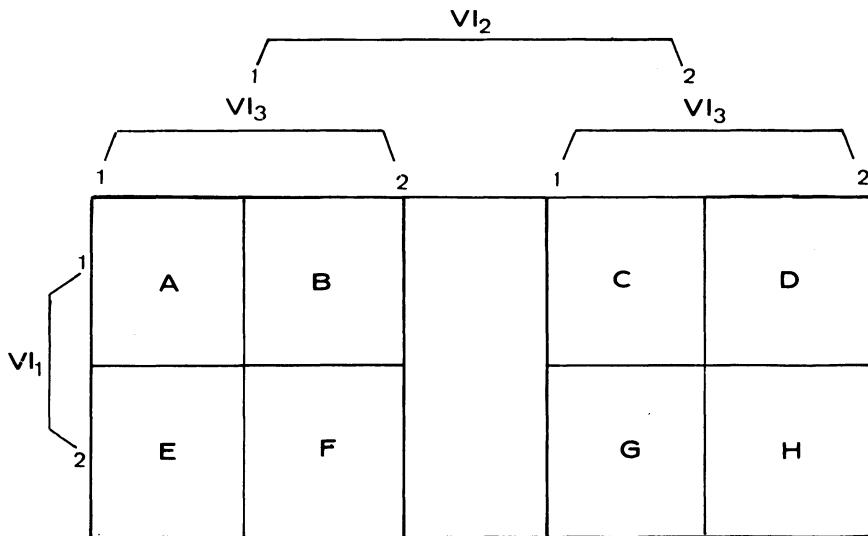
En el marco de la investigación educativa el área de medios y ayudas a la enseñanza, durante los últimos diez años, ha venido afrontando fuertes reformulaciones en sus modelos de referencia (Olson, D. 1974; Clark, R. 1983). Desde las etapas en que el medio era analizado a través de una serie de atribuciones específicas, lo que en muchos casos desembocaba en estudios comparativos de medios, paulatinamente se han ido aportando nuevos grupos de variables que han introducido constructos como: códigos de los medios, tareas de aprendizaje, expectativas de los sujetos, contextos de uso, entre otros. Esto, en definitiva está llevando a resituar los medios de enseñanza cara a su utilización en su contexto específico: el currículum (Escudero, J. M. 1983).

El enfoque dado a la investigación aquí reseñada parte de un modelo interactivo cuya base de análisis se encuentra en los diseños ATI (Aptitud y Tratamiento en Interacción), delimitados por Cronbach, L. y Snow, R. (1977). Nuestro modelo de análisis –ver gráfico 1–, se dimensiona en función de tres elementos: características del medio (cine); características del alumno usuario del medio y situación instruccional contextualizada (De Pablos, J. 1984). El trabajo experimental llevado a cabo intenta validar parcialmente dicho modelo.

La hipótesis principal específica que la interacción dada entre las características diferenciales de los alumnos, identificadas como dependencia e independencia de campo, y las características estructurales y temáticas de las películas didácticas monoconceptuales, debe tener una influencia significativa en el rendimiento y las actitudes de los sujetos de la muestra. Los factores interaccionantes controlados son tres: estilo cognitivo del alumno, tipo de estructura del film y contenidos. Han sido realizadas cuatro películas monoconceptuales (58 mm) en base al doble concepto de *alta y baja estructuración cinematográfica*. Entre una y otra existe una diferencia de grado en la presencia de los elementos sintácticos



**GRAFICO I.**  
MODELO PARA EL ANÁLISIS DEL CINE DIDÁCTICO



**GRAFICO II**

Las tres variables independientes aludidas son El Estilo cognitivo – D.I.C.– de los alumnos ( $VI_1$ ); la estructuración –alta o baja– del lenguaje de los films ( $VI_2$ ) y los temas presentados –«El acuario» y «Educación Vial» ( $VI_3$ ).

y semánticos del lenguaje filmico. En este caso hemos realizado dos versiones (alta y baja) de cada contenido.

El estilo cognitivo dependencia/independencia de campo (DIC) supone la expresión de la diferenciación de la estructura psicológica de la persona y es sin duda el más estudiado (Witkin, H. y Goodenough, P. 1985).

En estas breves páginas vamos a tratar de resumir exclusivamente algunos rasgos de la metodología de investigación aplicada.

El diseño experimental es factorial, de efecto fijo con tres factores, conteniendo cada uno de ellos dos niveles.

## LA MUESTRA

El método de muestreo utilizado ha sido probabilístico sustentado en la identificación de dos extractos: 1. Centros públicos de EGB de clase social medio-baja en Sevilla capital; 2. Centros privados de EGB de clase social medio-baja en Sevilla capital. (El factor de control de la procedencia social ha sido la profesión del cabeza de familia). Se ha aplicado la técnica de muestreo estratificado aleatorio con afijación proporcional. El tamaño de la muestra (5 centros, 475 alumnos de 12 y 13 años), se ha obtenido con una estimación de error de 0.05 y una probabilidad del 95% en base a la fórmula propuesta por Azorín, F. (1972, p. 337).

## INSTRUMENTOS DE EVALUACIÓN

Ha sido diseñada una *prueba cerrada* (prueba objetiva) compuesta por una serie de ítems de elección múltiple y aplicada en tres momentos distintos bajo las modalidades de pretest, postest y retest para cada uno de los dos contenidos seleccionados ("Los componentes de un acuario" y "Educación vial: giros"). Con ello se ha pretendido controlar algunas dimensiones de "reconocimiento" y "comprensión" (conocimiento) y de actitud hacia los temas.

También se ha diseñado una *prueba abierta* consistente en que cada alumno, inmediatamente después de la proyección del film, fue requerido para que formulase por escrito todas aquellas preguntas que la película le había sugerido. La categorización de tales preguntas cara a su codificación y análisis fue la siguiente:

- a) Número de preguntas realizadas.
- b) Número de preguntas sintácticas (códigos cinematográficos).
- c) Número de preguntas semánticas: convergentes, divergentes y no relacionadas (contenidos).

El "Group Embedded Figures Test" (Witkin, H. el al., 1971) ha sido la prueba utilizada para la identificación de los sujetos como dependientes o independientes de campo.

## TESTS ESTADÍSTICOS

El tratamiento estadístico se ha basado fundamentalmente en el análisis de covarianza (ANCOVA) para la prueba cerrada y el análisis de varianza (ANOVA) para la prueba abierta. Tanto en un caso como en otro se han controlado las condiciones paramétricas de normalidad, homoscedasticidad e independencia de los grupos experimentales.

## RESULTADOS DE LA INVESTIGACIÓN

Por la concisión exigida solamente vamos a hacer referencia a los resultados extraídos a partir de la prueba abierta.

De los tres factores identificados en el diseño experimental, en cuanto al número de preguntas formuladas, ha quedado contrastada de forma significativa la influencia del Estilo cognitivo ( $F = 11.51$ ) y del Tema del film ( $F = 10.47$ ), con un nivel de significación máximo (0.000 y 0.001). En cuanto a las interacciones, la acción conjunta estructura del film y tema ( $F = 55.71$ ) resulta significativa con un nivel de confianza máximo (0.000). Los independientes de campo resultan aportar mayor número de preguntas que los dependientes. Y el tema más influyente El Acuario.

En cuanto al número de preguntas semánticas-convergentes formuladas, es influida por la estructuración del film ( $F = 6.93$ ) y por el tema ( $F = 32.90$ ) con índices de confianza altos (0.009 y 0.000). Asimismo estos dos factores en interacción son influyentes ( $F = 5.95$  y 0.01). En consecuencia la formulación de preguntas convergentes sobre los contenidos de un film es influida por los factores mencionados, lo que implica que una determinada operación mental (producción convergente) pudiera ser suplantada, en determinados supuestos, por un film.

Finalmente, respecto a las preguntas semánticas-divergentes el análisis de varianza contrasta significativamente los factores Estilo cognitivo ( $F = 4.91$  y 0.02) y Tema ( $F = 26.21$  y 0.000) como ocurriría con la variable "número de preguntas formuladas", constatándose también la superioridad de los independientes de campo formulando preguntas divergentes y del tema El Acuario sobre el de Educación Vial.

En definitiva, tales constataciones deben suponer una referencia fiable a la hora de diseñar situaciones de instrucción en las que se utilicen los factores formulados en la investigación que ha servido de base para esta comunicación (De Pablos, J. 1984).

REFERENCIAS BIBLIOGRÁFICAS

- AZORIN, F.: (1972) *Técnicas de muestreo y sus aplicaciones*. Aguilar, Madrid.
- CLARK, R.: (1983) *Reconsidering Research on Learning from Media. Review of Educational Research*, V. 53, 4, pp. 445-459.
- CRONBACH, L. y SNOW, R.: (1977) "Aptitudes and Instructional Methods". Irvington Publish. Inc, New York.
- DE PABLOS, J.: (1984) "Propuesta de un modelo para el análisis del cine didáctico". *Cuestiones Pedagógicas*, N. 1, pp. 143-159.
- DE PABLOS, J.: (1984) *Variables estructurales del cine didáctico y su interacción con algunas características de los alumnos*. Tesis doctoral (inédita). Universidad de Sevilla.
- ESCUDERO, J. M.: (1983) "Nuevas reflexiones en torno a los medios para la enseñanza". *Revista Investigación Educativa*, N. 1, pp. 19-44.
- OLSON, D. (ed.): (1974) *Media and Symbols: The forms of expression, communication, and education*. University of Chicago Press.
- WITKIN, H. et al.: (1971) *A Manual for the Embedded Figures Test*. Consulting Psychologist Press, Palo Alto, California.
- WITKIN, H. y GOODENOUGH, D.: (1985) *Estilos cognitivos. Naturaleza y orígenes*. Pirámide, Madrid.

# Propuesta de un nuevo modelo de selectividad basado en estudios experimentales previos

por  
J. T. Pastor

## INTRODUCCIÓN

Todas las investigaciones experimentales acerca de la fiabilidad del modelo actual de pruebas de selectividad han puesto de manifiesto su inadecuación a tal fin:

“La prueba se selectividad, sin promediar con el expediente secundario explica nada más que el 6% del rendimiento posterior”. (Escudero, T. (1984) p-162). En particular, en lo que respecta a las matemáticas de C.O.U. las conclusiones fueron (Goberna, M. A. López, M. A. y Pastor, J. T. (1985).

“La importancia relativa de la prueba de ensayo frente a la prueba objetiva (40:60 para el total de la muestra frente a 55:45 para la muestra de aprobados) indica claramente que la influencia de las pruebas de acceso a nivel de ponderación de objetivos es, prácticamente, nula.”

Por otro lado y en un estudio en el que se trata de explicar el éxito en las Matemáticas del Primer Curso Universitario –en distintas facultades– a través de la NOTA de COU, (en Matemáticas) y de la NOTA de SELECTIVIDAD (Goberna, M. A., López, M. A., Pastor, J. T. y Valdivia, M. (1984)) se obtuvieron las siguientes conclusiones, basadas fundamentalmente en el recurso a modelos de regresión lineal (Achen, Ch. H. (1982) y paquete BMDP (1981)) y análisis de la varianza (Iversen, G. R. y Norpoth, H. (1976)).

- 1.º) Nivel predictivo muy bajo de la nota de selectividad ( $R^2 = 0.05$ ).
- 2.º) Nivel predictivo bajo de la nota de C.O.U. ( $R^2 = 0.10$ ).
- 3.º) No obstante, en los modelos de regresión múltiple, tan sólo la nota-COU resulta significativamente explicativa.
- 4.º) La influencia del tipo de educación recibida (estatal o pública) en el modelo de regresión es debida al azar con una probabilidad del 82%.

“Si los Centros Universitarios desearan conocer las aptitudes de los aspirantes a ingresar en las mismas... no cabe otra solución que establecer unas pruebas específicas (más adecuadas que las actuales) para las materias que cada centro considere fundamentales.”

## PROPUESTA DE UN NUEVO MODELO DE PRUEBAS DE SELECTIVIDAD

El nuevo modelo, que juzgamos viable, desde el punto de vista organizativo y económico, presenta, frente al modelo actual la ventaja de su individualización. Con otras palabras, la asignación, a cada alumno, de una puntuación que permita predecir, con la máxima fiabilidad, su rendimiento académico en el centro universitario al que aspira incorporarse. Procede comenzar describiendo nuestro modelo cuya concreción –experimental– debe ser objeto de un futuro proyecto.

### Principios generales

1. El objetivo es asignarle al aspirante una calificación que correlacione fuertemente con el futuro éxito en los estudios.
2. Para que el objetivo anterior sea realizable, no puede pensarse en establecer una prueba única para todos los Centros Universitarios: el modelo deberá tener en cuenta las peculiaridades de los diversos centros.
3. Sin perjuicio de lo dicho, deberá satisfacer al alumno unos mínimos culturales, que son irrenunciables de la condición universitaria.

### Las pruebas (descripción)

#### A. Prueba común

De cultura general, con calificación “apto” o “no apto”. Podría consistir en: conferencia, comentario de texto, test objetivo de conocimientos culturales (examen multi-respuesta) y (quizá) comprensión de un idioma extranjero.

#### B. Prueba específica

El aspirante realizaría los exámenes de las *materias* requeridas por la(s) Facultad(es) en la(s) que desea ingresar. La Universidad constituiría un único tribunal para cada una de estas materias. En el momento de solicitar el ingreso, el alumno debería consignar las calificaciones obtenidas en cada una de aquellas materias; de acuerdo con el baremo establecido para dicho centro y que debería hacerse público, el aspirante obtendría una puntuación indicativa de su expectativa de éxito, que constituiría el criterio único de selección. La convocatoria de los exámenes específicos podría hacerse simultáneamente en varias poblaciones del Distrito. En todo caso, el control debería corresponder a un tribunal único por materia. Quedaría así garantizada la competencia del tribunal y la aplicación de un criterio uniforme de corrección (uno de los aspectos más criticados del modelo actual).

## Comentario

- Enumeramos algunas de las ventajas del modelo propuesto frente al actual.
1. Se conseguiría (de acuerdo con los resultados de diversos trabajos empíricos) una notable correlación entre la calificación, según el baremo del Centro, y el futuro éxito en los estudios (correlación prácticamente nula en la actualidad).
  2. El sesgo que actualmente se produce, a causa de los diferentes criterios de los tribunales (injusticia inevitablemente asociada al modelo actual), desaparecería con los tribunales únicos por materias.
  3. No se correría el riesgo, como en la actualidad, de que el alumno fuera examinado, en parte de las materias, por profesores no especialistas.
  4. El modelo propuesto supone un ahorro de esfuerzos inútiles: un aspirante a la Facultad de Ciencias (Sec. Químicas) se debería examinar, probablemente, tan sólo, de Matemáticas, Física, Química; en la prueba específica para Ciencias Económicas sólo intervendrían, quizá, Matemáticas e Historia. Lo mismo ocurriría con los restantes centros.

## Diseño del baremo de cada Centro

Antes de seguir adelante, permítasenos justificar la necesidad de elaborar un baremo específico para cada Centro Universitario. Hemos probado (Goberna, M. A., López, M. A., Pastor, J. T. y Valdivia, M. (1984)) el escaso valor predictivo de un modelo único de regresión que explicara el “éxito” en las matemáticas universitarias a través de las calificaciones obtenidas en el centro de procedencia. Los resultados mejoraban notablemente cuando se consideraba la muestra procedente de cada centro universitario por separado.

Para diseñar cada baremo se requiere precisar las variables a considerar. La variable independiente “éxito en los estudios universitarios” se obtendrá a partir de dos índices: el primero, referido a la convocatoria de exámenes realizados, en junio, por el alumno durante su primer curso universitario. El segundo índice medirá el eventual estancamiento del alumno. Las variables independientes deberán medir la capacidad alcanzada por el alumno para cada una de las asignaturas que componen el plan de estudios vigente en los estudios preuniversitarios. Hasta que no transcurran tres años desde la implantación del nuevo modelo de selectividad será imposible relacionar las capacidades medidas a través de las pruebas específicas con la variable “éxito”.

La solución provisional no puede ser otra que tomar en consideración la calificación obtenida por el alumno, para cada asignatura en el centro de procedencia. Para justificar la eficiencia de tales variables permítasenos apelar a nuestra experiencia: hemos comprobado (Goberna, M. A., López, M. A., Pastor, J. T. y Valdivia, M. (1984)) que dichas calificaciones son aceptablemente fiables tanto si el centro es estatal como si es privado. Cada baremo inicial deberá ser modificado al término de los tres años de plazo arriba indicados y, posteriormente, con la periodicidad que se estime oportuno.

REFERENCIAS BIBLIOGRÁFICAS

- ACHEN, Ch. H. (1982): *Interpreting and Using Regression*, Sage University Papers.
- BMDP (1981): *BMDP Statistical Software*, 1981, University of California Press.
- ESCUDERO, T. (1984): "Condicionantes, capacidad predictiva de la Selectividad Universitaria". *Revista de Educación* 273, 139-164.
- GOBERNA, M. A.; LOPEZ, M. A.; PASTOR, J. T. y VALDIVIA, M. (1984): "Las Matemáticas en la transición de la Enseñanza Media a la Universidad", *Revista de Educación* 273, 49-63.
- GOBERNA, M. A.; LOPEZ, M. A.; PASTOR, J. T. (1985): "La influencia del examen de Selectividad en la enseñanza (Análisis de una experiencia en Matemáticas de COU)". *Enseñanza de las Ciencias*, (en prensa).
- IVERSEN, G. R. y NORPOTH, H. (1976): *Analysis of Variance*, Sage University Papers.

# Análisis multivariado causal de los trastornos de aprendizaje desde el diagnóstico clínico. Una aplicación del modelo log-lineal.

por  
*Maria Dolores Peris*

## INTRODUCCIÓN

Abordamos el tema, muy frecuente, del diagnóstico de los trastornos de aprendizaje durante los primeros años escolares. Los criterios y procesos de su evaluación son muy diversos, y para obtener cierta homogeneidad, acudimos a los Registros de los Servicios de Diagnóstico del INSERSO, donde un mismo equipo diagnostica a los niños con problemas escolares susceptibles de recibir subvención si se llega al diagnóstico de Deficiencia Mental (C.I. < 45). Seleccionamos los historiales con C.I. > 70, para eliminar de la muestra los trastornos asociados a las limitaciones en el desarrollo cognitivo. Los diagnósticos, a excepción de la edad y el C.I., estaban registrados en escalas nominales y en formato sin estructurar.

La estructuración de estas variables se realizó tomando como variable cada una de las afirmaciones y señalando en la matriz de datos la presencia (1) o ausencia (0) de cada variable en el informe de cada sujeto. La edad se dividió en 5 intervalos mútuamente exclusivos (uno para cada uno de los años desde 6 hasta 10), de modo que se registraba la «presencia» en la edad correspondiente al sujeto, y del mismo modo se dividió el C.I. en 4 intervalos (> 100, 90-99, 80-89, < 79). Las variables que exigían más de dos categorías, se dividieron en tantas variables como respuestas posibles (asi, hijo primogénito, hijo menor; parto con forceps, cesárea,...). Se acumuló un total de 57 variables fundamentalmente de dos clases: anamnesis y diagnósticos.

El análisis de estos datos dicotómicos ha sido tema de estudios en los últimos años, dada su importancia en las Ciencias sociales (Goodman, L.A. 1978; Reynolds, H.T. 1977; Feinberg, S.E. 1983; Knoke, D. y Burke, P.J. 1977; Nigel, G. 1981, entre otros). Los modelos loglineales aparecen aparecen en tales estudios como la alternativa más adecuada para llegar a establecer las relaciones entre variables, determinando las interacciones entre ellas, de forma equivalente

al modelo de ANOVA. En caso de que exista una relación temporal antes-después entre las variables analizadas, la causalidad puede ser evidenciada, en base a las relaciones entre las frecuencias en las tablas de contingencia. El análisis factorial de Bool, es otra alternativa para llegar a definir la estructura de las tablas del tipo «presencia-ausencia». Uno y otro análisis fueron realizados sobre la matriz de datos en que se registraron los historiales de 276 casos. Veamos los resultados de cada uno de ellos.

## ANÁLISIS FACTORIAL DE BOOL

Obtenemos los siguientes resultados:

Factor I: varón; C.I. > 100; 6 años; DISLALIA

Factor II: varón; DISLEXIA; COMPRENSIÓN DISMINUIDA; ATENCIÓN DISMINUIDA; MEMORIA DISMINUIDA.

Factor III: varón; 80 < C.I. < 90; COMPRENSIÓN DISMINUIDA.

Factor IV: lugar intermedio entre hermanos; DISLEXIA.

Factor V: C.I. < 79.

Los diagnósticos que aparecen en los factores, se corresponden con las variables que acumulan las mayores frecuencias en las características de «presencia». La dislexia (var. 15) se diagnostica en el 62% de los casos; la dislalia (var. 20) en el 36%, la disminución de la comprensión (var. 28), de la atención (var. 29) y de la memoria (var. 30) en el 68%, 58% y 52% respectivamente. El número superior de varones, como viene siendo habitual en estos estudios (66%), justifica su presencia en los tres primeros factores, si bien no se asocia con el último, referido a los cocientes más bajos. La dislexia aparece dividida en dos tipos: uno asociada con déficits de las actividades intelectuales (comprensión, atención y memoria), aunque no se asocia con ningún nivel concreto de C.I., y otro asociado con el lugar intermedio entre los hermanos, que generalmente se define, como el que menos atención recibe por parte de los padres. Si bien ambos tipos son independientes del C.I., parece estar vinculado el uno, con otros déficits cognitivos y el otro, con una posible falta de atención por parte de la familia.

En general, los tradicionales trastornos de aprendizaje diagnosticados (discalculia, disgrafía, disortografía,...) así como las variables causales más comunes, alcanzan tasas muy bajas, no presentando apenas asociaciones. Es de destacar que los valores del C.I., aun no apareciendo relacionados con trastornos concretos del aprendizaje, se ordenan por orden decreciente en tres de los factores obtenidos. Podría interpretarse este resultado como que la graduación del C.I. propuesta es la mejor clasificación, o diferenciación, del conjunto de variables analizadas. El primer factor, puede ser interpretado como un «pseudo-retraso», apareciendo los trastornos del lenguaje en los diagnósticos de los primeros años (var. 6), con un C.I. superior al medio (var. 2) y desapareciendo en años posteriores.

Tabla 1. – Resultados del Análisis Log-lineal. (Significación de efectos).

\*\*\*\*\* THE RESULTS OF FITTING ALL K-FACTOR MARGINALS.  
 THIS IS A SIMULTANEOUS TEST THAT ALL K-1 AND HIGNER FACTOR  
 INTERACTIONS ARE ZERO.

K-FACTOR	D.F.	LR CHISQ	PROB.	PEARSON CHISQ	PROB.	ITERATION
0-MEAN	255	614,11	.00000	1310,69	.00000	
1	247	265,12	.20435	422,53	.00000	2
2	219	86,70	1.00000	94,73	1.00000	7
3	163	45,13	1.00000	44,91	1.00000	3
4	93	15,10	1.00000	15,29	1.00000	4
5	37	3,99	1.00000	4,11	1.00000	4
6	9	,85	.99971	,86	.99970	3
7	1	,02	.87706	,02	.87706	2
8	0	,	1.	0,	1.	

\*\*\*\*\* A SIMULTANEOUS TEST THAT ALL K-FACTOR INTERACTIONS ARE  
 SIMULTANOUSLY ZERO.  
 THE CHI-SQUARES ARE DIFFERENCES IN THE ADOVE TABLE.

K-FACTOR	D.F.	LR CHISQ	PROB.	PEARSON CHISQ	PROB.
1	8	348,98	.00000	888,16	.00000
2	28	178,42	.00000	327,80	.00000
3	56	41,57	.92476	49,82	.70637
4	70	30,03	.99999	29,62	.99999
5	56	11,11	1.00000	11,18	1.00000
6	28	3,14	1.00000	3,25	1.00000
7	8	,83	.99911	,84	.99909
8	1	,02	.87706	,02	.87706

\*\*\*\*\* ASSOCIATION OPTION SELECTED FOR ALL TERMS OF ORDER LESS  
 THAN OR EQUAL TO 8 PARTIAL ASSOCIATION MARGINAL ASSOCIATION

EFFECT	D.F.	CHISQUARE	PROB.	ITER	CHISQUARE	PROB.	ITER
1	1	15,15	.0001				
2,	1	105,70	.0000				
3,	1	22,07	.0000				
4,	1	68,54	.0000				
5,	1	10,18	.0014				
6,	1	14,38	.0001				
7,	1	88,95	.0000				
8,	1	24,01	.0000				
12,	1	,46	.4980	7	,58	.4477	2
13,	1	,89	.3449	7	,90	.3433	2
14,	1	,24	.6209	7	,12	.7259	2
15,	1	,24	.6243	7	,89	.3453	2
16,	1	5,60	.0179	7	5,37	.0205	2
17,	1	,30	.5831	7	,21	.6447	2
18,	1	,04	.8409	7	,22	.6355	2
23,	1	,25	.6191	7	2,59	.1076	2
24,	1	23,24	.0000	4	40,28	.0000	2
25,	1	,78	.3779	6	4,85	.0277	2
26,	1	,44	.5048	6	0,54	.1052	2
27,	1	4,69	.0303	4	12,03	.0005	2
28,	1	14,36	.0002	4	25,96	.0000	2
34,	1	2,68	.1018	7	5,16	.0231	2
35,	1	,20	.6521	7	,77	.3809	2
36,	1	,07	.7867	7	,67	.4134	2
37,	1	5,12	.0236	7	2,01	.1563	2
38,	1	9,36	.0022	7	11,19	.0008	2
45,	1	4,58	.0323	7	13,16	.0003	2
46,	1	8,82	.0030	6	20,03	.0000	2
47,	1	3,09	.0788	4	9,25	.0023	2
48,	1	1,80	.1798	4	12,00	.0005	2
56,	1	22,59	.0000	7	30,81	.0000	2
57,	1	,00	.9717	7	,88	.3491	2
58,	1	,16	.6935	7	,56	.4526	2
67,	1	1,36	.2441	7	3,95	.0468	2
68,	1	,68	.4093	7	3,43	.0638	2
78,	1	3,76	.0525	6	7,22	.0072	2

## MODELO LOG-LINEAL.

Seleccionadas las 8 variables que presentan asociaciones más altas en los análisis de Kendal y de Spearman y cuyas frecuencias son más elevadas, realizamos con todas ellas una tabla de contingencia distribuyendo las frecuencias entre todas las combinaciones posibles de «presencia-presencia», «presencia-ausencia», «ausencia-presencia» y «ausencia-ausencia».

A continuación se estimó el logaritmo de la frecuencia esperada para cada celda, como una función aditiva de los logaritmos de los efectos e interacciones más elevadas, según el modelo linear no saturado.

tabla

En la tabla 1 aparecen sucesivamente los siguientes resultados:

- A) Prueba de significación de los efectos más altos para cada modelo.
- B) Prueba de significación de la suma de efectos para cada modelo.
- C) Prueba de significación parcial y marginal de cada uno de los efectos.

Las variables analizadas han sido:

1: varón;	2: C.I.>100	3: C.I. < 79	4: 6 años
5: DISLEXIA	6: DISLALIA	7: HIPERKINESIA	8: Comprensión disminuida.

Tanto en A) como en B) obtenemos que los modelos dejan de ser significativos a partir de las asociaciones de orden 3. Tan solo los efectos simples y algunas de las asociaciones entre dos variables son significativas. Por ello en C) tan sólo analizamos las asociaciones de orden 2 y las de las variables independientemente.

El modelo log-lineal no saturado (con sólo los efectos significativos) presenta las asociaciones señaladas en al fig. 1., según la ecuación:

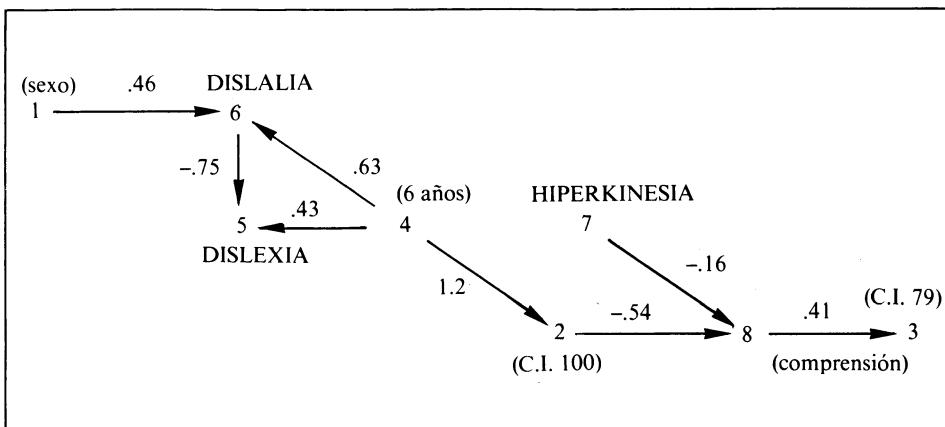
$$\theta + \lambda^1 + \lambda^2 + \lambda^3 + \lambda^4 + \lambda^5 + \lambda^6 + \lambda^7 + \lambda^8 + \lambda^{16} + \lambda^{24} + \lambda^{28} + \lambda^{38} + \lambda^{45} + \lambda^{46} + \lambda^{56} + \lambda^{78}$$

Este modelo, elegido como el mejor explica las relaciones entre las variables, permite establecer un diagrama causal dirigido las relaciones, en base a dos características de las variables:

a) Por ser las variables atributivas inmodificables y en consecuencia, origen y no término de la relación. Aplicando el modelo logit serían variables independientes que actuarían sobre las dependientes.

b) Por la distribución de las frecuencias entre las categorías de las variables, que muestran cuando la presencia en una variable está modulada por la presencia (relación positiva) o ausencia (relación negativa) en la otra variable.

Fig. 1. Modelo log-lineal



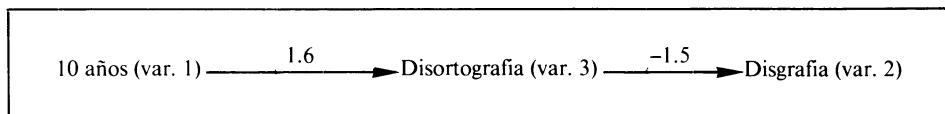
Algunas conclusiones a partir de este modelo son las siguientes:

1. La ausencia de Hiperkinesia aumenta el diagnóstico de Comprensión disminuida.
2. La presencia de Dislexia se ve especialmente favorecida por la ausencia de dislalia.
3. La relación entre Comprensión disminuida y C.I. < 79 es, como cabría esperar, positiva. La relación entre Hiperkinesia y C.I. > 100, es, sin embargo, positiva.
4. Los 6 años parecen decisivos para los diagnósticos que aquí se estudian, diferenciando dos grupos distintos de problemas: lenguaje y C.I.

Un segundo análisis log-lineal fue realizado sobre las variables disgrafía, disortografía y «10 años». Los resultados, como se muestran en la tabla 2, indican una significación equivalente, tanto en los efectos parciales como en los marginales, para las relaciones 13 y 23. El mejor modelo es el expuesto en la figura 2, según la ecuación

$$\theta + \lambda^1 + \lambda^2 + \lambda^3 + \lambda^{13} + \lambda^{23}$$

Fig. 2.- Modelo para las variables analizadas.



Obsérvese cómo en las tablas de frecuencias se evidencia el aumento del diagnóstico de disgrafía en ausencia del de disortografía, que a su vez aparece

Tabla 2. Resultados del análisis log-lineal. (Busqueda del mejor modelo).

\*\*\*\*\* OBSERVED FREQUENCY TABLE

3M 2M 1M  
AUSENCIA PRESENCI TOTAL

AUSENCIA	AUSENCIA	208	52	260
	PRES ENCI	7	2	
	TOTAL	215	54	269
PRES ENCI	AUSENCIA	1	3	4
	PRES ENCI	0	3	3
	TOTAL	1	6	7

TOTAL OF THE OBSERVED FREQUENCY TABLE IS

ALL CASES HAD COMPLETE DATA FOR THIS TABLE

\*\*\*\*\* DELTAS .500 IS ADDED TO EACH CELL FOR ALL ANALYSES

#### \*\*\*\*\* THE RESULTS OF FITTING ALL K-FACTOR MARCINALS.

**THIS IS A SIMULTANEOUS TEST THAT ALL K-1 AND HIGHER FACTOR INTERACTIONS ARE ZERO.**

K-FACTOR	D.F.	LR. CHISQ	PROB.	PEARSON CHISQ	PROB.	ITERATION
D-HEAN	7	704.51	.00000	1043.37	.00000	
1	4	26.85	.00002	123.11	.00000	2
2	1	,19	.66500	,18	.67141	5
3	0	0,	1.	0,	1.	

\*\*\*\*\* A SIMULTANEOUS TEST THAT ALL K-FACTOR INTERACTIONS ARE SIMULTANEOUSLY ZERO.  
THE CHI-SQUARES ARE DIFFERENCES IN THE ABOVE TABLE.

K-FACTOR	D.F.	LR CHISQ	PROB.	PEARSON CHISQ	PROB.
1	3	677,66	.00000	920,27	.00000
2	3	26,67	.00001	122,93	.00000
3	1	,19	.66500	,18	.67141

\*\*\*\*\* ASSOCIATION OPTION SELECTED FOR ALL TERMS OF ORDER LESS THAN OR EQUAL TO 3  
PARTIAL ASSOCIATION MARGINAL ASSOCIATION

EFFECT	D.F.	CHISQUARE	PROB.	ITER.	CHISQUARE	PROB.	ITER.
1,	1	92.06	.0000				
2,	1	276.99	.0000				
3,	1	308.58	.0000				
12,	1	,36	.5488	2	3,16	.0754	2
13,	1	10,32	.0013	2	13,12	.0003	2
23,	1	10,38	.0013	2	13,19	.0003	2
123,	1	,19	.6650				

\*\*\*\*\* ALL MODELS ARE REQUESTED

MODEL	D.F.	LIKELIHOOD. BATIO CHISQ.	PROB.	PEARSON CHISQ.	PROB.	ITERATIONS
1,	6	612.43	.0000	698.54	.0000	1
2,	6	427.52	.0000	436.81	.0000	1
3,	6	395.93	.0000	415.26	.0000	1
1,2	5	335.43	.0000	256.15	.0000	1
2,3	5	118.94	.0000	147.30	.0000	1
3,1	5	303.85	.0000	249.98	.0000	1
1,2,3,	4	26.85	.0000	123.11	.0000	1
12,	4	332.27	.0000	253.21	.0000	1
13,	4	290.73	.0000	233.00	.0000	1
23,	4	105.75	.0000	98.79	.0000	1
1,23,	3	13.67	.0034	17.22	.0006	1
2,13,	3	13.73	.0033	32.61	.0000	1
3,12,	3	23.69	.0000	64.71	.0000	1
12,13,	2	10.57	.0051	17.28	.0002	1
13,23,	2	,65	.7607	,54	.7650	1
23,12,	2	10.51	.0052	11.90	.0026	1
12,13,23,	1	,19	.6650	,18	.6714	5

después de los 10 años. Y cómo no puede mantenerse el modelo de independencia.

Algunos de los resultados obtenidos podrían plantear la hipótesis del «efecto del diagnóstico alternativo», de modo que cuando el sujeto carece de los síntomas de un diagnóstico determinado, aumenta la probabilidad de que se le dé un diagnóstico alternativo, apareciendo así las parejas de diagnósticos asociados con signo negativo, a la manera de constructos (Kelly) alternativos del diagnosticador.

#### REFERENCIAS BIBLIOGRAFÍCAS

- COX, D.R. (1970).: *Analysis of Binary Data*. Chapman and Hall. London.
- EVERIT, B.S. (1977).: *The Analysis of Contingency Tables*. Chapman and Hall. London.
- FIENBERG, S.E. (1983).: *The Analysis of Cross-Classified Categorical Data*. The MIT Press. Cambridge.
- GOODMAN, L.A. (1978).: *Analyzing Qualitative/Categorical Data*. Addison-Wesley. London.
- KNOKE, D. y BURKE, P.J. (1977).: *Log-linear Models*. Sage. Beverly Hills.
- MC CULLAGH, P. y NELDER, J.A. (1983).: *Generalized Linear Models*. Chapman and Hall. London.
- NIGEL GILBERT, G. (1981).: *Modelling Society*. G. Allen & Unwin. London.
- PLACKETT, R.L. (1981).: *The Analysis of Categorical Data*. Griffin & C. London.
- REYNOLDS, H.T. (1977).: *Analysis of nominal data*. Sage. Beverly Hills.

## Análisis diagnóstico de una realidad escolar Estudio experimental aplicado en E.G.B.

por

*Antonio de Pro  
Sebastián Sánchez*

### OBJETIVOS DEL ESTUDIO

Tras la aplicación de los contenidos programáticos de los Ciclos Inicial y Medio de la E.G.B., se hace necesario un análisis diagnóstico del marco real en el que se desenvuelve el proceso de enseñanza-aprendizaje (profesorado, características psicopedagógicas, interacciones antropológico-ambientales, modelos didácticos subyacentes,...). El hecho resulta de gran importancia, puesto que estos niveles educativos actúan de condición inicial de aprendizajes posteriores, perfilan la actitud del alumno hacia éstos, sientan las bases de un estilo docente y organizativo de las materias del currículo,... y, por supuesto, llevan implícitas unas repercusiones en el Área de las Ciencias Experimentales, donde se enmarca nuestro estudio.

Aunque son muchos los aspectos del mismo, nos vamos a centrar de forma aislada en dos puntos importantes desde la perspectiva cognitiva: la adquisición y la transferencia del aprendizaje en alumnos escolarizados, como elementos fundamentales del desarrollo de los mismos. Nos referimos, por limitaciones obvias, a una experiencia concreta, con unas técnicas específicas de análisis, una finalidad determinada y dentro de un proyecto de investigación en el que trabajamos en la actualidad.

### CARACTERÍSTICAS DEL CONTEXTO EDUCATIVO

Para caracterizar el contexto hemos utilizado, en primer lugar, las creencias del profesorado en ejercicio respecto a su acción educativa (programa, metodología, evaluación, autoconcepto, etc.), y unas variables de extracción social del alumnado (edad, sexo, profesión de los padres, colegios, etc.). Disponemos de la información correspondiente a 3.º, 5.º y 7.º de E.G.B. de Melilla (PRO, A. 1984;

Sánchez, S. 1985; Pro, A. y Sánchez, S. en prensa -a-), pero sólo exponemos aquí la del último curso, puesto que consideramos más interesante profundizar en las posibilidades de las pruebas utilizadas.

Los maestros entrevistados corresponden a Colegios Públicos (71% del total) y Privados (44% del total). Hay que resaltar la crítica que ellos hacen a los programas que ponen en práctica (en especial, falta de adecuación a la realidad local y excesivos contenidos), la escasa utilización de estrategias didácticas alternativas, la falta de conexión entre la vida del aula y los intereses del niño,... y el hecho de que en un 69% de los casos consideran su enseñanza eficaz.

Por otro lado, la muestra de alumnos la constituyan 369 alumnos de 7.<sup>º</sup> de E.G.B., de un total de 751 escolarizados, lo que supera el mínimo señalado por R.V. Krejcie y D.W. Morgan (Cf. Rivas, F. 1982) para el muestreo necesario en este estudio. Se escolarizan en seis colegios: cinco públicos y uno privado, tres de nueva creación y tres antiguos; dos ubicados en una zona de residencia de tipo medio-baja, tres de tipo medio y uno de tipo alto; cuatro con buena dotación de material y dos con menos medios.

## INSTRUMENTOS DE RECOGIDA DE DATOS

Hemos usado, además de los que nos han servido para caracterizar el contexto educativo (Pro, A. 1984; Sánchez, S. 1985), dos tipos de pruebas: experienciales y de destrezas básicas de laboratorio.

Con las pruebas experienciales pretendemos explorar la actitud del alumno ante problemas y situaciones reales, en los que debía transferir una serie de conocimientos relacionados con conceptos físicos fundamentales. Con ellas hemos procurado abarcar los contenidos más representativos del currículo escolar previo del niño en el Área que nos ocupa. En total eran siete y en ellas subyacían los conceptos de longitud, tiempo, velocidad, masa, fuerza, energía, calor, temperatura, carga eléctrica y propiedad de reflexión de la luz. No obstante, más importante que esto, era que en las situaciones exploratorias se daban procesos de descripción de observaciones, lógico-deductivos, de interpretaciones, de identificación conceptual, de coherencia en razonamientos de predicciones, de valoraciones de situaciones,... y el hecho de que por sus características pueden resultar de bastante utilidad al profesorado en ejercicio, dadas sus posibilidades de adaptación a realidades concretas (Para una información más detallada Cf. Pro, A. 1984; Sánchez, S. 1985; Fernández, A., Pro, A. y Sánchez, S. en prensa; Pro, A. Sánchez, S. y Fernández, A. en prensa).

En cuanto a las pruebas de destrezas básicas de laboratorio, fueron diseñadas con la finalidad de explorar otro factor fundamental para el aprendizaje de la Física, la experimentación. Se trataba de estudiar el grado de eficacia de los alumnos para la manipulación inteligente de diversos instrumentos y materiales de laboratorio. Hemos buscado actividades en las que los sujetos debían utilizar ciertos aparatos físicos, aplicar algunos principios experimentales básicos, elegir

estrategias para resolver problemas prácticos, reconocer conceptos en situaciones reales, etc.; siendo la principal diferencia con las anteriores la necesidad de manipulación por parte del alumno. Es decir, se pretende estudiar el comportamiento de los sujetos en el laboratorio. La selección de instrumentos, montajes y aparatos se hizo en función de lo que los programas oficiales contemplan en los niveles más bajos del sistema educativo. Entre ellos, la regla, el transportador, el cronómetro, el dinamómetro, el termómetro, la balanza, etc., que, como se puede apreciar, responden a las exigencias señaladas. En la exploración se planteaba el problema a resolver, además de indagarse sobre la utilidad del instrumento, el proceso manipulativo seguido, las estrategias usadas en la resolución, la identificación conceptual, las valoraciones perceptivas, etc. (Pro, A. 1984; Pro, A. y Sánchez, S. en prensa -b-).

## **RESULTADOS OBTENIDOS CON LAS PRUEBAS DE ESTRATEGIAS BÁSICAS DE LABORATORIO**

Cuando se trabaja con una muestra grande y con instrumentos de evaluación no habituales en las situaciones de aprendizaje escolar, aparecen situaciones anecdóticas y una casuística enormemente rica para el análisis y discusión de los resultados. Sin embargo, por razones de espacio, sólo nos referimos a algunos datos que den una idea descriptiva de la exploración efectuada. En primer lugar ofrecemos una síntesis numérica de las respuestas obtenidas en la tabla n.º 1.

**TABLA I**

	X	$\sigma_n$	$X_W$	%	Prueba experimental
L.1	2.61	1.56	4	65.22 %	Regla
L.2	1.32	1.38	4	32.88 %	Cronómetro
L.3	0.61	1.04	4	15.29 %	Dinamómetro
L.4	0.52	0.95	4	12.91 %	Probeta
L.5	0.67	0.87	4	16.85 %	Balanza
L.6	0.61	1.12	4	15.35 %	Termómetro
L.7	0.45	1.06	4	11.28 %	Volumen de un sólido
L.8	1.13	1.61	4	28.33 %	Transportador
L.9	2.23	1.46	4	55.64 %	Conexión circuito
L.10	0.43	0.88	4	10.87 %	Amperímetro

Como puede apreciarse, a pesar de que todos los problemas planteados deberían haber sido abordados en los niveles anteriores del sistema educativo, sólo dos pruebas pasan de un rendimiento medio (superior al 50%). En concreto se trata del «uso de la regla» y de la conexión de elementos de un circuito. El resto obtienen porcentajes de acierto inferiores, estando en seis pruebas por debajo del 20%.

Para completar algo más estos resultados, destacamos los siguientes análisis parciales:

- Hay 169 alumnos (46%) que han cometido errores importantes en el uso de la *regla*.
- 194 alumnos (53%) no supieron decir lo que marcaba el *cronómetro*, y otros 93 (25%) sólo dieron respuestas aproximadas al valor real.
- Cabe destacar que 248 alumnos (67%) no conocían el *dinamómetro*; además 169 (46%) ni siquiera han acertado las unidades que venían especificadas en el propio aparato.
- Señalamos que 312 (85%) no conocían el uso que tiene la *probeta* (205 la perciben sólo como un recipiente para echar agua). Análogamente al caso anterior, 264 (72%) vuelven a errar en las unidades especificadas en el recipiente.
- Sólo el 6% indica que la *balanza* sirve para medir la masa, dándose bastantes casos de errores manipulativos y de balances intuitivos disparatados.
- En el caso del *termómetro*, 81 alumnos (22%) presentan confusiones entre calor y temperatura; 317 (86%) tienen errores en la lectura y 97 (26%) fallan al señalar O.<sup>o</sup> y/o 100.<sup>o</sup> C en el instrumento, a pesar de que viene especificado en el mismo.
- El *volumen del sólido* es acertado por 86 alumnos (23%), repartiéndose entre ambos procedimientos, regla y Principio de Arquímedes, en proporción equivalente.
- Hay 190 alumnos (52%) que cometen errores en la utilización del *transportador de ángulos*.
- Nos ha llamado la atención, sobre todo en comparación con los resultados anteriores, que 297 alumnos (80%) *conectan los elementos del circuito* en el tiempo estipulado.
- Sólo 81 alumnos (22%) conocen el nombre del *amperímetro*, confundiéndolo, en otros casos, con cronómetros, termómetros, etc., perteneciendo la mayoría de éstos (66) a un mismo colegio.

A la vista de los resultados, es difícil entender su incoherencia con el nivel educativo que ocupan los sujetos, puesto que los errores van más allá de dificultades experimentales, propias de la utilización de material de laboratorio en situaciones reales.

Un primer análisis de los resultados por Clases (dos por cada colegio de la muestra), nos indica una mayor heterogeneidad que en las pruebas experienciales (PRO, A. 1984; Sánchez, S. 1985) entre los colectivos, incluso dentro del mismo Centro. En concreto, realizados los contrastes estadísticos correspondientes, se obtienen diferencias significativas entre los grupos A y los B de cinco colegios (LS, RC, GV, ES y AM). Sólo en uno existe equivalencia estadística.

En los contrastes realizados entre Centros (encaminados a analizar los efectos de variables no controladas), también se observan diferencias significativas entre los Centros AM y RC y todos los demás. Los rendimientos oscilan entre un 34% para AM y un 19% de ES, siempre inferiores a los que presumiblemente corresponden a los alumnos de 7.<sup>º</sup> de E.G.B. Podría interpretarse que la influencia del profesor es mayor que la del propio sistema escolar, pues existe una omisión generalizada del uso de estrategias básicas experimentales dentro de un marco concreto de aprendizaje, aunque los programas oficiales indiquen lo contrario. Sin embargo, contrastan estos valores con el 69% de este mismo profesorado que consideraba su enseñanza como eficaz. Este desconocimiento de la situación diagnóstica de los alumnos es, sin duda, una de las limitaciones más importantes en cualquier alternativa metodológica o didáctica, en el caso de que éstas se pretendan.

## ANÁLISIS FACTORIAL DE LAS PRUEBAS DE ESTRATEGIAS BÁSICAS DE LABORATORIO

Para estudiar más detenidamente los distintos problemas experimentales utilizados, procedimos a la aplicación de la técnica del análisis factorial por el procedimiento VARIMAX. El programa usado es el BMD P4M de J. Dixon (1977).

Los resultados de la matriz de correlaciones rotada se reflejan en la tabla n.<sup>º</sup> 2. Los tres factores explican el 43.60% de la varianza. Sin embargo, a pesar de que en las pruebas experienciales existía una «lógica» en la factorización encontrada (Pro, A., Sánchez, S. y Fernández, A. en prensa), no ocurre así en este caso, aún tomando el nivel de saturación en 0.45. Quizás el factor III (balanza y termómetro) tenga su justificación en el alto poder discriminador de las dicotomías conceptuales peso-masa y calor-temperatura, que subyacen en las pruebas.

TABLA II

	Factor 1	Factor 2	Factor 3
Amperím.	0.739	—	—
Transport.	0.643	—	—
Vol. sólid.	0.626	—	—
Dinám.	—	0.709	—
Con. Circ.	—	0.571	0.309
Cronómet.	—	0.547	—
Balanza	—	—	0.747
Termómetro	—	0.416	0.593
Probeta	0.472	—	0.281
Regla	0.346	—	—

Respecto a los otros factores, es difícil encontrar una interpretación de la agrupación resultante, aunque tal vez respondan a la proximidad al pensamiento concreto del desarrollo evolutivo del alumno. Esto precisaría de nuevos estudios o de una mayor profundización en los resultados encontrados. Para ello, hemos realizado un análisis de las relaciones existentes entre los valores obtenidos con las pruebas de estrategias básicas de laboratorio y con las experienciales.

### RELACIONES ENTRE LOS DOS TIPOS DE PRUEBAS

El procedimiento estadístico utilizado ha sido el análisis de la matriz de correlaciones correspondientes a ambos tipos de prueba y con posterioridad el análisis factorial, realizado mediante el programa BMD P4M, antes citado.

Han resultado siete factores, que explican el 52.40% de la varianza. Una representación de la estructura simple aparece en la tabla n.º 3, según los siguientes criterios:

x: Saturación 0.400 – 0.499  
 xxx: Saturación 0.600 – 0.699

xx: Saturación 0.500 – 0.599  
 xxxx: Saturación 0.700 – 0.800

TABLA III

TIPO	REF.	F.1	F.2	F.3	F.4	F.5	F.6	F.7
C	L			xxxx				
O	t							
N	v			xxx				
C	m							
E	f							
P	E							
T	T							
O	Q							
S	q	xx						
	RL			xx				
L	Reg.			x				
A	Cro.				xxx			
B	Din.				xxx			
O	Prb.	xx						
R	Bal.							
A	Ter.							
T	Vol.	xx						
O	Ang.	xx						
R	Cir.				x			
I	Amp.	xxxx						
O								

Se observa, en primer lugar, que en cuatro de los factores (F. 4, F. 5, F. 6 y F. 7) se agrupan sólo conceptos o experimentos, pero sin solaparse entre sí. ¿No contradicen estos resultados la idea de una mayor dependencia entre los procesos de transferencia conceptual y experimental en el aprendizaje de la Física?. Por otro lado, mientras existe una coherencia en la factorización de las unidades conceptuales (F. 2 agrupa a conceptos cinemáticos, F. 4 los de mecánica, F. 5 los de termodinámica,...), no resulta esta misma «lógica» en las de laboratorio. ¿Es acaso el aprendizaje conceptual más estructurado que el experimental?.

Con todo, no encontramos una interpretación inmediata de F.1 y F. 3, que precisamente reunen las declaraciones de los factores 1 y 2 del análisis anterior. El problema es importante, puesto que explican el 21.80% del total de la varianza y contienen nueve de las veinte variables con las que hemos trabajado. Ello nos llevará a nuevos estudios empíricos, no sólo para resolver la problemática factorial, sino para tratar de responder algunos de los interrogantes anteriormente apuntados.

#### REFERENCIAS BIBLIOGRÁFICAS

- DIXON, J. (1977): *Biomedical Computer Programs BMD*. University California Press, Los Angeles.
- FERNÁNDEZ, A.; PRO, A. y SÁNCHEZ, S. (en prensa): «Adquisición de conceptos físicos por el niño en un sistema escolar «monolateralizado». Ponencia al IV Seminario sobre Investigaciones actuales en Psicología y Educación. I.C.E. de la Universidad Autónoma de Madrid.
- PRO, A. (1984): *Diseño, aplicación y evaluación de dos metodologías didácticas (expositiva-audiovisual y experimental-integrada) para la enseñanza de la Física en un nivel elemental*. Tesis doctoral inédita. Facultad de Física. Universidad de Sevilla.
- PRO, A.; SÁNCHEZ, S. (en prensa -a): «Aplicación de aprendizajes conceptuales a situaciones reales. Un estudio sobre los conceptos de longitud, tiempo y velocidad». *Enseñanza de las ciencias*.
- (en prensa -b): «Un estudio exploratorio sobre la actividad experimental del niño en E.G.B.». *Publicaciones*. Universidad de Granada.
- PRO, A.; SÁNCHEZ, S. y FERNÁNDEZ, A. (en prensa): «Adquisición de estrategias manipulativas y de conceptos físicos en niños de E.G.B. de Melilla». Trabajo presentado al IV Encuentro de Didáctica de Física y Química. I.C.E. de la Universidad de Cádiz.
- RIVAS, F. (1982): *Elementos de Psicometría (I)*. Edic. Rubio Estaban. Valencia.
- SÁNCHEZ, S. (1985): *Transfer y generalización de conceptos físicos a situaciones de experiencia (descontextualización curricular). (Población E.G.B.)*. Tesis doctoral inédita. Facultad de Filosofía y Ciencias de la Educación. Universidad de Valencia.

# Dimensiones de un modelo para la investigación de la expresión escrita infantil

por  
Carlos Rosales

## ¿EN QUÉ CONSISTE LA ESCRITURA?

La investigación didáctica, si bien concede en la actualidad un creciente interés a los lenguajes paraverbales (entonación, ritmo...) y no verbales (utilización del espacio, movimientos del cuerpo, las extremidades, los ojos...), no ha dejado en el olvido el estudio del uso de la palabra, del lenguaje verbal, dado su papel de protagonista en el conjunto de intercambios comunicativos y situaciones de expresión personal.

La escritura constituye una de las habilidades básicas implícitas en la utilización del lenguaje verbal. En ella se acumulan las mayores dificultades pues a las inherentes a las habilidades anteriores hay que añadir las derivadas del trazado de los signos gráficos, así como del carácter mucho más convencional y restringido de su utilización.

Existe estrecha relación entre el aprendizaje y uso oral de la palabra y su utilización gráfica a nivel mental y simbólico, dándose en este último caso una doble posibilidad correspondiente a la relación inmediata de los signos con sus correspondientes significados o bien, con la utilización intermedia de los signos orales correspondientes. Se manifiesta la estrecha relación entre las dimensiones oral y gráfica de la palabra especialmente en los estados patológicos del uso del idioma. Las distintas modalidades convencionalmente diferenciadas de dislexia oral visual y gráfica, presentan importantes repercusiones en los restantes ámbitos de la comunicación verbal.

La metodología didáctica, antaño centrada en la ejercitación de los aspectos más formales de la expresión escrita, tiende, de forma progresiva, desde el movimiento de Escuela Nueva, al desarrollo de las habilidades implícitas en la realización de la misma: motivación para la comunicación, desarrollo de la capacidad de organización espacio-temporal, de secuencialización, de la oculomotricidad y la grafomotricidad...

## DIMENSIONES PARA SU INVESTIGACIÓN.

En estrecha relación con las características específicas de la expresión escrita, parece conveniente que el modelo de investigación de la misma reuna una serie de dimensiones como las siguientes: omnicomprensiva, genética, criterial y contextual.

### Omnicomprensiva

Investigar la naturaleza de la expresión escrita significa profundizar en el conocimiento de todos sus componentes. Una visión completa tendría que partir de la consideración de aspectos como:

—Motivación /—Organización mental /—Simbolización /—Trazado y utilización de los signos gráficos.

Un modelo para la investigación de la expresión escrita debería basarse en el estudio en profundidad del conjunto total de características que la componen para el establecimiento de juicios fundamentados y la elaboración de diagnósticos sobre las causas de sus posibles problemas, en el caso de darse. Podemos pensar, por ejemplo, que en un determinado problema de aprendizaje pueden incidir cuestiones relativas a la falta de maduración, a la carencia de contenidos, a la incapacidad para la organización coherente del discurso. Dentro de cada uno de estos factores pueden identificarse a su vez importantes subcomponentes. Así, en el aprendizaje de la ortografía concurren tres tipos de imágenes del signo escrito: auditiva, visual y cinética. La imagen auditiva interviene en el dominio de la denominada ortografía natural, la imagen visual en el de la ortografía convencional y en ambos casos la imagen cinética constituye un apoyo importante.

Otro ejemplo representativo lo puede constituir el trazado de los signos gráficos, en el que se dan, según Ajuriaguerra, autor de la escala "E", aproximadamente treinta rasgos dependientes en mayor o menor medida de la evolución genética o la ejercitación. Un ejemplo más abreviado de instrumento para el diagnóstico de la capacidad del alumno para la escritura podría lograrse simplificando el modelo anterior en otro que contuviera diez rasgos como los siguientes (Rosales, C, 1984):

- 1)–Calidad en el trazado de las curvas
- 2)–Calidad en los trazos rectos
- 3)–Continuidad en el trazado de las letras
- 4)–Destreza en el trazado de las mayúsculas
- 5)–Retoques
- 6)–Unión de letras
- 7)–Tamaño y regularidad de zonas
- 8)–Progresión de la escritura
- 9)–Líneas
- 10)–Inclinación de las letras

Podríamos referirnos asimismo al vocabulario, con matices relativos a su nivel de comprensión y utilización en la expresión, su nivel de concrección / abstracción, etc.

### Genética

La capacidad de comunicación escrita se desarrolla con la edad. Estudios analítico-descriptivos sobre la expresión gráfica infantil ponen de relieve que con la edad se produce:

- \* Un incremento de la capacidad mental para la recuperación, organización y expresión de contenidos.
- \* Un incremento de la capacidad de elaboración de estructuras discursivas.
- \* Un incremento de la longitud y complejidad sintáctica de la expresión.
- \* Un incremento del vocabulario común y específico.
- \* Un desarrollo progresivo de la capacidad para la realización de trazos de calidad.
- \* Un desarrollo progresivo de la velocidad de escritura.
- \* Una creciente intensificación del uso del idioma como instrumento para la comunicación, (Piaget, J., 1976) y para la actividad mental interiorizada (Vygotsky, L., 1977).

Como manifestación específica del carácter genético de la evolución de la expresión escrita se pueden citar los estudios de J. de Ajuriaguerra (1977) sobre adquisición de la capacidad grafomotriz, en la que intervienen, según este autor, los componentes del tono, fuerza, localización y rapidez. El tono como capacidad para separar la cabeza de la escritura o el cuerpo de la mesa así como para el mantenimiento de un equilibrio dinámico entre cuerpo y extremidades en el acto de la escritura. La fuerza se manifiesta tanto en la realización más perfecta de los movimientos de inscripción como en el mismo hecho de coger el útil de escritura. La localización del movimiento varía con la edad desde una primera fase de gran difusión hasta la utilización de determinados centros de apoyo, primero en el hombro y posteriormente en el codo, muñeca y mano. De entre las diversas conclusiones a que llega Ajuriaguerra en sus estudios sobre grafomotricidad, se pueden citar los siguientes:

- a) En el periodo comprendido entre los seis y los once años se registran dos fases de progreso rápido (seis a ocho y diez a once años) y uno de estabilización y consolidación de los trazos (ocho a diez años).
- b) Las niñas presentan una cierta ventaja sobre los niños en la calidad de los trazos, que se manifiesta en la aproximación de la media de las niñas al cuartil superior de los niños y temporalmente en una ventaja aproximada de seis meses.
- c) La velocidad en la escritura evoluciona asimismo con la edad y especialmente

en los períodos de seis a ocho años y de diez a once tanto en situación de escritura a ritmo normal como acelerado. En este aspecto no se aprecian diferencias significativas entre niños y niñas.

### **Criteria**

La investigación del lenguaje y, específicamente, de la expresión escrita, bajo supuestos criteriales más que normativos, implica la consideración de los distintos códigos o lenguajes como cualitativamente distintos mas que como jerárquicamente superiores o inferiores. A partir de los primeros estudios de Bernstein se llegó a considerar la existencia de formas de expresión distintas e implícitamente, más o menos perfectas. Dada la utilización por la escuela de la forma más acabada, se deducía la necesidad de "compensar", de dotar a los alumnos menos favorecidos de una serie de aprendizajes tempranos que los colocasen en igualdad de condiciones con otros sectores más favorecidos respecto al nivel "normal" del lenguaje utilizado en el medio escolar. La práctica de una enseñanza, como de una investigación de tipo "normativo" acaba perjudicando en último término a aquellos grupos que más se distancian de dicho nivel o norma. Por ello parece pedógicamente más adecuado partir del estudio de las características de la expresión lingüística del grupo, tratando de acomodar las características de la enseñanza e investigación a la naturaleza específica de las mismas.

### **Contextual**

La conveniencia de que la investigación del lenguaje sea de carácter contextual puede entenderse en diversas manifestaciones. Se puede entender, en principio, como coordinación de los distintos componentes de la expresión escrita. Así, el estudio del vocabulario no puede realizarse con independencia del contexto sintáctico y semántico en que aparecen los términos. Mezynski, K. (1983) ha puesto de relieve cómo el conocimiento del significado de los términos que componen un texto a veces resulta innecesario y otras veces insuficiente para la comprensión global del mismo.

Otra dimensión de la investigación contextualizadora de la expresión escrita se centra en la necesidad de considerarla como una forma de expresión más entre las varias de posible aplicación por la persona en un momento determinado. La expresión escrita constituye un instrumento más de comunicación entre los diversos lenguajes a utilizar. Presenta unas características determinadas y resulta especialmente útil en ciertas situaciones.

La tercera perspectiva de contextualización se centra en la necesidad de establecer vínculos entre la práctica escolar de la escritura y las formas de utilización de la misma en el ámbito social extraescolar.

REFERENCIAS BIBLIOGRÁFICAS

- AJUARIAGUERRA, J. (1977): *La escritura del niño*. Laia, Barcelona.
- BERNSTEIN, B. (1971, 73, 75): *Class, codes and control*. Vol, I, II, III Routledge and Kegan, London.
- BOUTON, CH. (1979): *El desarrollo del lenguaje*. Huemul/UNESCO. Buenos Aires.
- BRUNER, J. (1984): *Acción, pensamiento y lenguaje*. Alianza. Madrid.
- DALE-JORDAN, E. (1980): *La dislexia en el aula*. Paidós, Buenos Aires.
- LABOV, W. (1978): *Sociolinguistique*. Ediotions de Minuit, París.
- LABOV, W. (1978) : *Le parler ordinaire*. Editions de Minuit, París.
- MEZYNSKI, K. (1983): «Cuestiones relativas a la adquisición de conocimientos. Efectos del aprendizaje del vocabulario en la comprensión lectora». *Review of Educational Research*. Summer, 1983. vol. 53, N.º 2, pág. 253-279.
- PERRENOUD, PH. (1983): «Des differences culturelles aux inegalités scolaires: l'évaluation et la norme dans un enseignement indifférencié», en la obra de ALLAL, L. (1983): *L'évaluation formative dans un enseignement différencié*. Peter Lang, Berne.
- PIAGET, J. (1976): *Lenguaje y pensamiento en el niño*. Guadalupe, Buenos Aires.
- PLOWDEN RAPPORt (1967): Central Advisory Council for Education Children and their Primary Schools. London.
- ROSALES, C. (1984): *Didáctica de la comunicación verbal*. Narcea, Madrid.
- VYGOTSKY, L. (1977): *El pensamiento y el lenguaje*. La Pléyade, Buenos Aires.

# Influjo de las técnicas de estudio en el rendimiento académico del adolescente

por  
*Francisco de Salvador*

## INTRODUCCIÓN

El rendimiento escolar es un producto complejo, en el que inciden múltiples factores, estrechamente relacionados. Esta investigación analiza algunos factores, referidos al discente: inteligencia, adaptación psicosociológica y hábitos de estudio.

Algunos investigadores han encontrado relaciones positivas entre inteligencia y rendimiento; más altas, incluso, que con otros aspectos de la personalidad. En cambio, otros investigadores no encuentran relaciones significativas (Pérez Serrano, G. 1981, págs. 265-280). Algunos autores han obtenido correlaciones significativas entre rendimiento y otras dimensiones básicas de la personalidad.

Pocos estudios han analizado la influencia de los hábitos y técnicas de estudio (Rogero, A. 1975; Baquero, E. 1977; Ossola, E. y Horas, P. 1977; Gallardo, M.D. 1978; Salvador, Mata, F. 1984). Todos encontraron relaciones positivas con el rendimiento académico.

La presente investigación tiene dos objetivos: 1) determinar si el rendimiento académico de los adolescentes se relaciona con algunos aspectos de su personalidad (inteligencia, adaptación psicosocial y hábitos de estudio); 2) determinar el peso específico de cada uno de los aspectos mencionados, en especial los hábitos de estudio.

La experiencia cotidiana de muchos profesores indica que el rendimiento académico depende en gran medida, aunque no exclusivamente, de los hábitos y técnicas de estudio. Esta investigación pretende validar científicamente esta intuición.

## METODOLOGÍA DE LA INVESTIGACIÓN

### Hipótesis

H-1: «Los hábitos de estudio del adolescente se relacionan positivamente con el rendimiento académico en más alto grado que otras dimensiones de su personalidad (inteligencia y adaptación psicosocial)».

H-2: «Algunos aspectos de los «hábitos de estudio» se relacionan con el rendimiento académico en más alto grado que otros».

### Muestra

Constituida por 72 sujetos, alumnos de un curso terminal de Bachillerato, en un centro privado de Valencia. Su edad estaba comprendida entre los 15 y 17 años. Todos los estratos de clase social estaban representados.

### Variables e instrumentos de medida

1. *Variable dependiente*: «rendimiento académico», medido por la calificación final del curso (global). La escala verbal se tradujo a escala numérica: Insuficiente = 4; Suficiente = 5; Bien = 6; Notable = 7, Sobresaliente = 8; Matrícula de Honor = 9.

#### 2. *Variables independientes (8 variables)*

A. *Hábitos de estudio*: cuatro variables, correspondientes a las 4 escalas del «Cuestionario» de Brown, W.F y Holtzman, W.H. (1975), que mide:

- 1) «Diligencia en el estudio»: puntualidad y regularidad en las tareas escolares
- 2) «Método de trabajo»: utilización de procedimientos eficaces para el estudio y realización de tareas escolares.
- 3) «Actitud ante los profesores»: sobre su persona y su actuación en el aula.
- 4) «Aceptación de fines»: aceptación racional de la finalidad del estudio y sus requisitos.

B. *Inteligencia*, medida por el test de matrices progresivas, PM-62 (Raven, J.C., 1968).

C. *Adaptación psicosocial*: tres variables (1.<sup>a</sup> adaptación familiar; 2.<sup>a</sup> adaptación a la salud; 3.<sup>a</sup> adaptación social), medidas por un cuestionario (Bell, H.M. 1970). Se tomaron sólo tres de las cuatro escalas del cuestionario, por deficiencias en el expediente psicológico de los alumnos.

3. *Variables interviniéntes*: diferencias de edad, profesores, exámenes y calificaciones. Los profesores son los mismos para todo el grupo. La calificación representa un juicio, contrastado entre varios profesores, sobre diversas materias. Hay unos elementos homogeneizadores, más importantes que la edad: idéntico nivel instructivo (administración), mismo sistema de trabajo docente, parecida problemática psicológica de los alumnos.

### Técnicas estadísticas

Se aplicaron técnicas correlacionales. Para determinar la correlación del rendimiento académico (variable dependiente) con cada una de las variables independientes (Hipótesis-1) se aplicó la correlación lineal de Pearson. Para la comprobación de la hipótesis-2 se obtuvo una correlación múltiple, siguiendo el método de Doolitee (Yela, M., págs. 68-80).

Se aplicaron test de significación y fiabilidad a los diversos coeficientes de correlación. Se fijó un nivel de significación del .05.

## ANÁLISIS Y DISCUSIÓN DE RESULTADOS

### Rendimiento académico y factores de personalidad

Los coeficientes de correlación más altos corresponden a «los hábitos de estudio», como preveía la hipótesis-1. Se exceptúa la variable «actitud ante los profesores», que no resultó significativa. Por el contrario, los coeficientes de correlación de los otros aspectos de la personalidad estudiados son muy bajos y no significativos. Resulta sorprendente la escasa correlación entre inteligencia y rendimiento ( $r = .06$ ).

A pesar de su escasa relación con el «rendimiento» ( $r = .07$ ) la «adaptación familiar» correlaciona alto con los «hábitos de estudio»: «aceptación de fines» ( $r = .513$ ); «actitud ante los profesores» ( $r = .392$ ); «diligencia en el estudio» ( $r = .383$ ); «método de trabajo» ( $r = .354$ ). Cuando el sujeto se siente integrado en la familia, se esfuerza en sus tareas escolares y utiliza los procedimientos adecuados que le conduzcan al éxito escolar. Lo contrario sucede en tantos de nuestros alumnos adolescentes.

La «adaptación a la salud» sólo correlaciona significativamente con la «adaptación familiar» ( $r = .497$ ) y con la «aceptación de fines» ( $r = .295$ ), lo que significaría que quien se percibe físicamente fuerte, se propone y se esfuerza por conseguir metas ambiciosas.

### Hábitos de estudio y rendimiento académico

La correlación más alta con el criterio corresponde a la variable «diligencia en el estudio» ( $r = .404$ ). Teniendo en cuenta la escasa entre inteligencia y rendi-

miento, se confirma la opinión de muchos profesores: los alumnos que obtienen mejores calificaciones no son los más inteligentes sino los que más se esfuerzan.

Sigue en importancia la variable «método de trabajo», cuya correlación con el criterio es .363. Relación obvia, por cuanto el éxito escolar está en función de un método de trabajo adecuado. Hay que subrayar la relación, lógica por lo demás, entre «diligencia» y «método» ( $r = .473$ ). En efecto, quien se esfuerza por aprender, emplea un método adecuado para conseguirlo.

Igualmente alta la correlación entre «método» y «actitud ante los profesores» ( $r = .404$ ). En definitiva, el «método» aparece como un importante predictor del rendimiento académico.

«Aceptación de fines» correlaciona moderadamente con el criterio ( $r = .310$ ). En cambio, correlaciona alto con los otros factores de «hábitos de estudio»: «diligencia en el estudio» ( $r = .648$ , la más alta correlación econtrada en esta investigación); «actitud ante los profesores» (.628). Curioso: de todas las variables, «aceptación de fines» correlaciona en más alto grado con «inteligencia» (.204). Lógico, ya que los fines se conciben racionalmente.

«Actitud ante los profesores» no correlaciona significativamente con el criterio ( $r = .139$ ). Se puede tener buena calificación, aunque no se acepte al profesor (crisis de la adolescencia?).

Unidos los cuatro factores de «hábitos de estudio», su valor predictivo aumentó ligeramente sobre el factor más significativo («diligencia...»). El R fue de .455. Sólo el 20.72% de la varianza es explicado por los cuatro factores asociados. Sin duda, el resultado no es demasiado satisfactorio.

El factor «diligencia», sólo, explica el 16.32% de la varianza. Unido a los otros factores, el 11.83%. A pesar de ello, es el mejor predictor del rendimiento académico. El factor «método», aislado, explica el 12.96% de la varianza; unido a los otros factores, sólo el 8.49 %.

## CONCLUSIONES

En los términos establecidos en esta investigación: 1.<sup>a</sup>. Los «hábitos de estudio» se relacionan positivamente con el rendimiento académico del adolescente más que la inteligencia o la adaptación psicosocial (familiar, fisiológica, social).

2.<sup>a</sup> La «diligencia en el estudio» y el «método de trabajo» son los factores de los hábitos de estudio más relacionados con el rendimiento académico y, por ende, los que mejor predicen el éxito académico del adolescente.

3.<sup>a</sup> La inteligencia y la «adaptación psicosocial» del adolescente no parecen relacionados con su rendimiento académico. La adaptación «familiar» y «fisiológica» aparecen relacionados indirectamente, a través de los «hábitos de estudio».

4.<sup>a</sup> Si se quiere obtener un rendimiento óptimo del alumno adolescente, se debe fomentar la diligencia en el trabajo y la ordenación racional del tiempo dedicado al estudio.

REFERENCIAS BIBLIOGRÁFICAS

- BAQUERO, E. (1977): *Estudio experimental de algunas variables psicológicas, pedagógicas, somáticas y sociológicas que influyen en el rendimiento escolar*. Universidad Complutense, Madrid.
- BELL, H.M. (1970): *Cuestionario de adaptación para adolescentes*, Herder, Barcelona.
- BROWN, W.F. y HOLTZMAN, W.H. (1975): *Encuesta de hábitos y actitudes hacia el estudio*, Tri-llas, México.
- GALLARDO, M.D. (1978): «Afición lectora y su relación con el rendimiento escolar», *Revista de Ciencias de la Educación*, 96, pp. 607-25.
- OSSOLA, E. y HORAS, F. (1977): «Inteligencia, hábitos de estudio y rendimiento escolar en la adolescencia inicial», *Revista de la Instituto de Investigaciones Educativas*, III, 10, pp. 23-46.
- PÉREZ SERRANO, G. (1981): *Origen social y rendimiento escolar*, Centro de Investigaciones Sociológicas, Madrid.
- RAVEN, J.C. (1968): *Test de Matrices Pregresivas*, Paidós, Buenos Aires.
- ROGERO, A. (1975): *Infijo de las técnicas de estudio en el rendimiento académico*, Universidad Complutense, Madrid.
- SALVADOR MATA, F. (1984): «Hábitos de estudio y rendimiento escolar», *Escuela de Maestros*, 3, pp. 49-62.
- YELA, M. (s/f): *Psicometría y Estadística*, Escuela de Psicología y Psicotecnia, Madrid.

# Análisis de datos cualitativos en la investigación sobre el rendimiento académico: resolubilidad ante casillas nulas mediante modelos Quasi-log-linear

por  
*Pedro Sánchez Algarra*

Históricamente, en la mayor parte de los estudios sobre el rendimiento académico, se ha enfatizado el análisis de datos de medida, relegándose el tratamiento de los cualitativos a un mero análisis descriptivo. Cuando el estudio interesa ser llevado a cabo puntualmente (o en un seguimiento a partir de una sucesión más o menos prolongada de puntos de tiempo) y los datos de que se dispone son categóricos, el análisis «log-linear» es un útil instrumento que en los últimos años se ha expandido con fuerza, a pesar de que sus orígenes se situaran hacia 1960 debido, fundamentalmente, a los esfuerzos de Goodman, que culminaron casi dos décadas después (Goodman, L.A. 1978). De aquí que en aquellos casos en que se disponga de valoraciones cualitativas o informes acerca de diferentes aspectos a considerar sobre el rendimiento, basta tener en cuenta tantos criterios clasificatorios –y por tanto tantas categorizaciones o taxonomías dicotómicas o polítómicas– como dimensiones vayan a ser consideradas, y en base a las cuales se partirá de una tabla de contingencia a resolver.

La construcción y análisis de tabulaciones cruzadas es una de las estrategias que últimamente se han extendido en la investigación sobre el rendimiento académico, y en este sentido basta recordar las ricas posibilidades existentes, por ejemplo, en el análisis de tareas cognitivas (Rodrigo, M.J. 1982), o en cualquier situación en que no sea suficiente una recogida de datos cuantitativos para captar la riqueza que subyace en una determinada estrategia que hay que valorar. Partiendo de tales premisas, la finalidad de estas líneas es la de aportar algunas sugerencias que agrandan sus posibilidades en el caso atípico de casillas nulas.

En efecto, existen una serie de restricciones que deben cumplirse satisfactoriamente (Kennedy, J.J. 1983), y entre ellas se halla la ocupación de todas las casillas de la tabla de contingencia. Debe distinguirse, en las casillas vacías, entre ceros de muestreo y ceros estructurales.

Los primeros (Fienberg, S.E. 1977) no son inherentes a la situación estudiada, y dependen de la muestra global, no suponiendo más que la ausencia de casos representativos en la tabla de clasificaciones cruzadas; si no es muy elevado el número de casillas vacías, se puede solventar mediante la adición de una pequeña cantidad (0,5) al cero existente, con lo cual se logra la necesaria viabilidad para el cálculo de las frecuencias esperadas. Cuando los ceros de muestreo son numerosos, sin embargo, pueden surgir inconvenientes, dado que se carecería de la necesaria información para inferir adecuadamente, la potencia estadística sería escasa, y se vulnerarían además las suposiciones que son básicas para un adecuado uso de la  $\chi^2$ -cuadrado (teniendo en cuenta que el análisis log-linear implica una superación simultánea de  $\chi^2$ -cuadrado y ANOVA), cuya distribución, en muestras grandes, se aproxima a la ley multinomial.

Otro carácter presentan los ceros estructurales, que proceden de la ausencia de frecuencias en tablas en las que se sabía de antemano que era imposible la obtención de datos no nulos (Gilbert, 1981). La tabla resulta evidentemente incompleta, afectándose sus grados de libertad –que se reducen– y haciendo necesario su ajuste (Bishop, Y.M.M., Fienberg S.E. & Holland P.W., 1975), dando lugar a los modelos *quasi-log-linear* (Upton G.D.G., 1978), que estructuralmente son similares a los log-linear, ya que generan frecuencias esperadas para la independencia de forma habitual, salvo en las filas o columnas que contienen uno o más ceros.

En muchos casos, es difícil que los propios investigadores reconozcan el hecho de la existencia de ceros estructurales, y con elevadísima frecuencia se produce el abandono del estudio en el análisis de datos, o bien se llega a resultados e interpretaciones inadecuadas. Y, en el caso en que origina menos trastornos, se tiende a modificar el sistema de categorías.

En tablas bidimensionales, que es el caso más frecuente, si  $S$  consiste en una serie de casillas en una ordenación  $I \times J$  resultante después de eliminar las casillas vacías, si  $x_{ij}$  es la frecuencia observada en la casilla  $(i,j)$  y  $m_{ij}$  la correspondiente frecuencia esperada, consideramos que en las casillas que no se incluyen en la serie  $S$  es evidente que

$$x_{ij} = m_{ij} = 0, \quad (1)$$

de forma que podemos seguir utilizando la notación habitual para los marginales totales.

Así, si  $S$  se compone de todas las casillas excepto  $(1,1)$

$$m_{1+} = \sum_{j=2}^J m_{1j} = \sum_{j=1}^J m_{1j} \quad (2)$$

representa el valor esperado para la primera fila de la tabla incompleta  $S$ .

Al utilizar los mismos modelos «log-linear» para ésta, y por analogía con el ANOVA, para las casillas  $(i,j) \in S$ , tenemos

$$\log m_{ij} = u + u_{1(i)} + u_{2(j)} + u_{12(ij)} \quad (3)$$

donde

$$\sum_{i=1}^I u_{1(i)} = \sum_{j=1}^J u_{2(j)} = 0 \quad (4)$$

y

$$\sum_{i=1}^I \sigma_{ij} u_{12(ij)} = \sum_{j=1}^J \sigma_{ij} u_{12(ij)} = 0 \quad (5)$$

con

$$\sigma_{ij} = \begin{cases} 1 & \text{para } (i,j) \in S \\ 0 & \text{en los demás casos} \end{cases} \quad (6)$$

Los términos  $u_{12}$  de (5), que corresponden a casillas que no pertenecen a S son series iguales a una cantidad arbitraria finita, de forma que (5) está bien definido.

A partir de aquí, el modelo de cuasi-independencia se halla estableciendo

$$u_{12(ij)} = 0 \text{ para } (i,j) \in S \quad (7)$$

de forma que

$$\log m_{ij} = u + u_{1(i)} + u_{2(j)} \text{ para } (i,j) \in S \quad (8)$$

En otras palabras, las variables correspondientes a las filas y columnas son cuasi-independientes si podemos escribir  $\{m_{ij}\}$  en la forma:

$$\left\{ m_{ij} \right\} \text{ en la forma:} \\ m_{ij} = \begin{cases} a_i b_j & \text{para } (i,j) \in S \\ 0 & \text{en los demás casos} \end{cases} \quad (9)$$

La cuasi-independencia actúa, pues, como la independencia, y se aplica a las casillas no vacías de la tabla.

Por otra parte, la regla general para hallar los grados de libertad, se concreta (Fienberg S.E., 1981):

g.l. = número casillas – número parámetros ajustados (10)  
y aquí es igualmente aplicable. Si hay e casillas eliminadas, entonces la serie S contiene  $IJ-e$  casillas, y el número de parámetros ajustados son

$$I + J - 1 \text{ (1 por } u, I-1 \text{ por } u_{1(i)}, \text{ y } J-1 \text{ por } u_{2(j)})$$

resultando, en consecuencia,

$$IJ - e - (I + J - 1) = (I - 1).(J - 1) - e \text{ grados de libertad}$$

Es interesante considerar su repercusión práctica en estudios acerca del rendimiento académico, ya que en situaciones en donde una serie de casillas de la tabla son ocupadas por ceros es frecuente optar por modificar el sistema de categorías, con lo cual se pueden evitar tales ceros, pero, en cambio, se sacrifica capacidad discriminatoria entre tales categorías, hecho que puede llegar a ser especialmente preocupante si la escala es ordinal (Andrich, D. 1979), apareciendo, en cualquier caso, y como consecuencia, los efectos de errores de clasificar inadecuadamente (Fleiss, V.C. 1981).

Las posibilidades metodológicas de optimización son evidentes, y en el ámbito del rendimiento académico pueden llevarse a la práctica con relativa facilidad, por lo que la propuesta realizada de utilizar modelos quasi-log-linear ante la existencia de ceros estructurales es de esperar que produzca sus frutos.

#### REFERENCIAS BIBLIOGRÁFICAS

- ANDRICH, D. (1979) A model for contingency tables having an ordered response classification. *Biometrics*, 35, 403-415.
- BISHOP, Y.M.M.; FIENBERG, S.E. & HOLLAND, P.W. (1975) *Discrete multivariate analysis: Theory and practice*. Cambridge, Mass.: MIT Press.
- FIENBERG, S.E. (1981) *The analysis of cross-classified categorical data*. Cambridge, Mass.: MIT Press.
- FLEISS, J.L. (1981) *Statistical methods for rates and proportions*. New York: Wiley & Sons.
- GOODMAN, L.A. (1978) *Analyzing qualitative/categorical data*. Cambridge, Mass.: Abt Books.
- KENNEDY, J.J. (1983) *Analyzing qualitative data. Introductory log-liner analysis for behavioral research*. New York: Praeger.
- RODRIGO, M.J. (1982) Las posibilidades del análisis de tareas como técnicas para el estudio de los procesos mentales. *Infancia y Aprendizaje*, 19-20, 159-173.
- UPTON, G.J.G. (1978) *The analysis of cross-tabulated data*. New York: Wiley & Sons.

## «Pruebas objetivas y calificaciones escolares»

por  
*Nicolás Seisdedos*

Los datos estadísticos que, de manera resumida, ofrecemos aquí han sido elaborados a partir de las puntuaciones directas recogidas por el CIDE («Evaluación de la Reforma de las Enseñanzas Medias», aplicaciones de Octubre-Noviembre, 1985).

De la muestra total (más de 25.000 casos) de 1.º de BUP y 1.º de FP de centros situados en localidades rurales y urbanas de casi todas las provincias españolas, se seleccionaron los casos que tenían datos completos en las pruebas objetivas de rendimiento, en las evaluaciones escolares, en los tests de aptitudes y en otras variables de tipo contextual y socioeconómico. La muestra resultante se clasificó, para la presente comunicación, únicamente por el sexo (resultando 11.146 varones y 8.070 mujeres). Las variables, con su denominación, especificación y amplitud de la escala de medida, han sido:

<u>Nombre</u>	<u>Especificación</u>	<u>Escala</u>
---------------	-----------------------	---------------

### PRUEBAS (rendimiento objetivo)

1. ORTOG	Ortografía .....	0-44
2. MTCAL	Matemáticas-Cálculo .....	0-30
3. MTAPL	Matemáticas-Aplicaciones .....	0-15
4. COMLEC	Comprensión lectora (de A. Lázaro) .....	0-28

### EVALUACIONES (rendimiento subjetivo)

5. NOTLEN	Nota en Lengua .....	0-5
6. NOMAT	Nota en Matemáticas .....	0-5
7. NOCINA	Nota en Ciencias Naturales .....	0-5
8. NOCISO	Nota en Ciencias Sociales .....	0-5

---

9. NOTIEX	Nota en Idioma Extranjero .....	0-5
10. NOGLO	Nota en Global .....	0-5

### APTITUDES

11. PALDIF	TEA-3, Palabra Diferente .....	0-20
12. VOCAB	TEA-3, Vocabulario .....	0-30
13. FV	TEA-3, Factor verbal (PALDIF + VOCAB) .....	0-50
14. FR	TEA-3, Factor razonamiento abstracto .....	0-30
15. FN	TEA-3, Factor numérico .....	0-30
16. APTES	TEA-3, Total (FV + FR + FN) .....	0-110
17. ADESP	TEA-3, Aptitud espacial, Rotación Figuras Macizas	0-21
18. DAT-MR	TEA-3, Razonamiento mecánico .....	0-68

En la tabla 1 presentamos las intercorrelaciones ( $r_{xy}$  de Pearson) en milésimas y los estadísticos básicos ( $\bar{X}$  = media aritmética,  $s_x$  = desviación típica) de las 18 variables; En la mitad superior derecha los estadísticos se refieren a la muestra de varones, y en la mitad inferior izquierda a la muestra de mujeres. Se han separado con líneas los tres bloques de variables (pruebas, evaluaciones y aptitudes) para una mejor lectura de los índices.

Con unos grupos (N) tan grandes como estos, basta un índice de correlación muy pequeño (0,24 en varones y 0,029 en mujeres) para ser significativo al nivel de confianza del 1%. En la tabla todos los índices lo son al N.c. del 1%, es decir, los tres conjuntos y todas sus variables tienen algo en común. No obstante, para una mejor comprensión e interpretación de su significatividad convendría tener en cuenta la distinta variabilidad ( $s_x$ ) de los resultados. Además, en la comprensión de los estadísticos e índices de correlación se recordará que hay dos variables con varianza espúrea:

$$13 = 11 + 12$$

$$16 = 13 + 14 + 15 \text{ (o bien } 16 = 11 + 12 + 14 + 15\text{)}$$

El contenido de la tabla puede observarse/analizarse desde varios puntos de vista:

- diferencias, entre ambas muestras (varones y mujeres) y en variables de igual o distinta área de medida (rendimiento, evaluaciones y aptitudes).
- variabilidad que presentan las escalas en ambas muestras.
- cuantía, atendiendo al grado de significatividad de los estadísticos e índices presentados, y
- relaciones, de tipo simple ( $r_{xy}$ ), parcial ( $r_{y,z}$ ) y múltiple ( $R_{1,2345\dots}$ ).

En el aspecto diferencial, nuestro compañero M. Álvaro y colaboradores han presentado en esta reunión sus comentarios a éstas y otras variables exami-

TABLA I

Var.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	$\bar{X}$	sx		
1. ORTOG	466	403	374		346	272	277	281	346	370	322	297	354	340	386	438	138	156	33,62	7,56		
2. MTCAL	378	689	445		373	424	333	325	338	436	433	362	451	507	591	624	235	295	14,84	6,57		
3. MTAPL	328	653	454		365	424	337	322	338	434	436	369	457	481	616	624	277	342	4,85	2,94		
4. COMLEC	296	386	408		290	229	245	279	250	314	423	427	488	382	407	524	205	310	12,61	4,11		
5. NOTLEN	315	320	335	262		619	641	657	634	860	266	255	299	243	298	342	77	104	3,61	1,19		
6. NOMAT	257	407	413	224		629	645	525	528	808	236	181	235	286	349	349	143	165	3,65	1,27		
7. NOCINA	247	334	339	244		662	677	639	541	842	230	219	257	192	251	285	94	130	3,88	1,22		
8. NOCINA	247	304	316	265		670	574	679	531	815	252	259	284	182	246	296	43	68	3,92	1,25		
9. NOTIEX	308	318	314	226		659	566	584	569	788	250	234	277	239	284	325	85	101	3,68	1,27		
10. NOGLO	328	403	411	292		861	825	859	834	808	300	279	331	278	348	388	108	138	3,75	1,02		
11. PALDIF	252	365	392	391		243	230	230	247	216	278		507	827	474	451	731	232	288	6,76	2,86	
12. VOCAB	237	290	311	362		233	173	210	233	191	248		439	904	370	400	704	173	272	10,39	3,76	
13. FV	286	380	408	441		279	233	257	281	238	308		808	884		476	484	822	228	321	17,16	5,76
14. FR	278	442	316	343		175	252	169	144	176	219		404	319	419	550	828	315	339	12,60	5,36	
15. FN	298	532	542	357		246	323	227	204	251	299		388	329	418	493	800	269	318	12,28	4,47	
16. APTES	360	558	562	481		292	333	273	264	275	343		687	662	792	824	757	330	399	42,04	12,75	
17. APESP	120	215	238	183		64	106	62	43	62	81		208	130	194	309	268	322	419	7,52	484	
18. DAT-MR	135	287	311	285		118	153	124	103	102	144		289	275	331	350	318	421	315	41,72	7,59	
$\bar{X}$ $s_x$	36,63 15,66 5,95 4,37 2,68 13,73 3,83	4,01 1,18 3,84 1,26 4,10 1,21 1,25 4,12 1,25 4,03 1,03	6,60 2,69 9,60 3,39 16,20 5,16 5,30 10,56 3,81 40,32 11,33 4,78 3,79 34,90 6,87																			

nadas en el estudio general («Evaluación de la Reforma...»), aunque tenemos que aclarar que en este momento (Gijón, Sep. 1985) se hizo la salvedad de que los datos podrían tener algunas deficiencias de grabación, y los que presentamos en estas páginas ya están libres de dichas deficiencias. No obstante, son válidas todas las consideraciones presentadas por M. Álvaro.

En cuanto a los demás aspectos, y dada la riqueza de implicaciones que observamos en la tabla, vamos a intentar un doble enfoque: Primero, hacer por nuestra cuenta algunas observaciones y presentar los resultados de unos pocos análisis complementarios (esperando que el lector interesado se anime a hacer otros más significativos para él); segundo, esperar que éste observe y analice los datos y saque sus propias conclusiones. De modo general, queremos señalar los puntos siguientes:

- A. Todos los índices son positivos y significativos, así como las submatrices que quedan encuadradas y señalan las tres áreas de medida empleadas.
- B. En promedios, los varones son superiores en aptitudes, y las mujeres lo son en rendimiento (excepto en Matemáticas Aplicaciones) y en evaluaciones.
- C. Las mujeres son más homogéneas (menor  $s_x$ ) en casi todas las medidas y, consecuentemente, sus índices de correlación son más significativos.
- D. Excepto en la submatriz de las evaluaciones, los índices de relación son menores en la muestra de mujeres, lo cual parece apuntar a una mayor independencia o madurez de las variables de rendimiento y aptitudes.
- E. Las pruebas de rendimiento objetivo se relacionan poco con las evaluaciones, y no en exclusividad con la materia afín.
- F. Las intercorrelaciones de las evaluaciones son elevadas (entre 0,53 y 0,66) y no es de extrañar que las pruebas objetivas correlacionen casi por igual con las distintas evaluaciones. Hay bastante covarianza entre éstas y, examinando la tabla, hemos pensado que puede ser debido a los aspectos intelectuales.
- G. Las aptitudes (sobre todo el resumen que significa la variable 16. TEA-Total), presentan con las evaluaciones unas relaciones casi tan altas como las pruebas de rendimiento objetivo.
- H. Las aptitudes espacial y de razonamiento mecánico (var. 17 y 18) tienen muy poca relación con las evaluaciones, pero su valor predictivo de las pruebas de rendimiento es similar al de otras aptitudes. ¿Está interviniendo también en las pruebas de rendimiento una inteligencia general subyacente a todas las pruebas de aptitudes?

Y en este momento dejamos al lector la tarea de seguir anotando observaciones/conclusiones sobre la tabla 1.

Por nuestra parte hemos querido ahondar un poco más lo señalado en el punto E, sobre todo teniendo en cuenta lo indicado en los puntos F, G y H. En

**TABLA 2**  
**Matriz R de intercorrelaciones  $r_{xy.z}$  de Rendimiento y Evaluaciones**  
**cuando se ha eliminado la influencia de inteligencia.**

Var.	1	2	3	4	5	6	7	8	9	10
1. ORTOG		275 229	185 163	189 151	233 236	142 156	177 166	177 169	240 234	242 234
2. MTCAL		272 229	491 495	178 162	218 198	282 283	208 228	188 196	183 207	270 272
3. MTAPL		185 163	491 495	191 190	207 217	282 290	213 234	184 211	183 201	267 281
4. COMLEC		189 151	178 162	191 190	139 145	58 78	118 134	153 164	99 112	142 155
5. NOTLEN		233 236	218 198	207 217	139 145		568 590	604 633	620 643	589 630
6. NOMAT		142 156	282 283	282 290	58 78	568 590		608 647	472 535	468 524
7. NOCINA		177 166	208 228	213 234	118 134	604 633	608 647		606 655	495 551
8. NOCISO		177 169	188 196	184 211	153 164	620 643	472 535	606 655		482 536
9. NOTIEX		240 234	183 207	183 201	99 112	589 630	468 524	495 551	482 536	760 791
10. NOGLO		242 234	270 272	267 281	142 155	840 847	779 803	828 847	796 821	760 791

**TABLA 3**  
**Relación múltiple entre las evaluaciones y las cuatro**  
**pruebas de rendimiento**

Eval.	R	%	Ecuación de regresión (Betas) en «z»								
NOGLO	0,356	12,71	0,164	z1 +	0,137	z2 +	0,159	z3 +	0,056	z4	
	0,368	13,56	0,163	z1 +	0,137	z2 +	0,172	z3 +	0,075	z4	
NOTLEN	0,308	9,50	0,171	z1 +	0,104	z2 +	0,111	z3 +	0,067	z4	
	0,315	9,93	0,184	z1 +	0,077	z2 +	0,134	z3 +	0,079	z4	
NOMAT	0,332	11,04	0,063	z1 +	0,176	z2 +	0,188	z3 -	0,021	z4	
	0,342	11,69	0,086	z1 +	0,618	z2 +	0,193	z3 +	0,001	z4	
NOCINA	0,275	7,58	0,114	z1 +	0,103	z2 +	0,131	z3 +	0,053	z4	
	0,297	8,81	0,104	z1 +	0,122	z2 +	0,143	z3 +	0,071	z4	
NOCISO	0,265	7,05	0,116	z1 +	0,090	z2 +	0,100	z3 +	0,096	z4	
	0,285	8,11	0,111	z1 +	0,090	z2 +	0,128	z3 +	0,108	z4	
NOTIEX	0,287	8,23	0,195	z1 +	0,072	z2 +	0,106	z3 +	0,029	z4	
	0,186	9,21	0,186	z1 +	0,102	z2 +	0,112	z3 +	0,096	z4	

primer lugar hemos tomado la submatriz de las diez primeras variables (rendimiento y evaluaciones), y hemos calculado los correspondientes índices cuando de las interrelaciones ( $r_{xy}$ ) se elimina la influencia de la inteligencia (var. 16. TEA-Total); es decir, hemos obtenido la matriz  $r_{xy,z}$  de las citadas diez variables.

En la tabla 2 se recogen los resultados. Como las relaciones con la inteligencia son positivas (véase la tabla 1), todos los índices de la submatriz 2 han descendido en comparación con los de 1. En cada casilla de cruces de variables, el primer índice se refiere a la  $r_{xy,z}$  en la muestra de varones y el segundo a la de mujeres. Se observa que todavía siguen altos los índices de la submatriz de las evaluaciones, y siguen siendo significativas las relaciones entre rendimiento y evaluaciones. ¿Continuaría habiendo las relaciones entre las evaluaciones si en distintos pasos se eliminan las influencias de los conocimientos o pruebas de rendimiento, es decir, si se calculan los índices  $r_{xy,1234}$ ? Probablemente quedaría todavía varianza común debida a variables de la personalidad del alumno y de la del evaluador, o contextuales, o socioeconómicas.

Por el momento, nos hemos centrado en otro enfoque. A partir de los datos de la tabla 2 hemos calculado las correlaciones múltiples de cada evaluación con las cuatro pruebas de rendimiento, para conocer la varianza común entre cada rendimiento objetivo y las distintas evaluaciones. En la tabla 3 se presentan dichas R múltiple, el porcentaje de varianza común y la ecuación de regresión (con los coeficientes Beta) en puntuaciones típicas «z»; la primera línea se refiere a la muestra de varones y la segunda a la de mujeres.

# El modelo de desamparo aprendido aplicado al análisis del rendimiento: principios conceptuales y metodológicos y recursos técnicos.

por  
Felipe Trillo

## INTRODUCCIÓN

Esta comunicación no aporta conclusiones. Lo que pretende es presentar en sus principios conceptuales, recursos técnicos y alternativas de análisis estadístico, un posible modelo de análisis para una de las manifestaciones del rendimiento académico: el fracaso escolar.

En concreto, este modelo está basado en la reciente importancia de las investigaciones psicosociales que desde un «marco de referencia interno» se preocupan de la significación y repercusión que tienen determinadas variables en el individuo, *pero tal y como él las percibe* (Rodríguez Espinar S, 1982, p. 62).

Pues bien, por cuanto adoptar esta línea de investigación pretendiendo mantenerse dentro de un estricto marco de referencia interno no se agota con el estudio del autoconcepto, sino que exige estar alerta a las características que conforman el peculiar sistema de creencias de los sujetos, así como de los prejuicios que hayan podido ir acumulando en el baúl de los recuerdos, es por lo que el *Modelo de Desamparo Aprendido* nos ha parecido muy sugerente en cuanto a sus posibilidades de ayudar a la comprensión del fracaso escolar.

A nuestro juicio, su interés radica no sólo en que nos permita diagnosticar que el sistema de enseñanza en general sitúa al alumno en una serie de situaciones que podrían ser descritas en términos de estricta «incontrolabilidad» (concepto que, como veremos, es fundamental en la investigación sobre «desamparo aprendido»); sino en que, nos ha permitido plantear la hipótesis de que *los reiterados fracasos escolares conducen al sujeto a una situación de desamparo aprendido*; esto es, que la acumulación de evidencias de fracaso (el suspender evaluación tras evaluación, el repetir curso,...) debe tener repercusiones en determinados parámetros individuales como, por ejemplo, pueden ser: la interrupción de proyectos de logro, o la ausencia de control sobre el entorno próximo.

## PRINCIPIOS CONCEPTUALES

En el Modelo de Desamparo Aprendido es posible distinguir dos momentos: uno, el que partiendo de la experimentación con animales, dió origen a todo el entramado teórico sobre el «desamparo» y que fue denominada Teoría del Desamparo Aprendido (Seligman, M.E.P., 1975/1981; y dos, el que tras numerosos estudios (Wortman, C.B. y Brehm, J., 1975) que advertían de la inconsistencia de esta formulación para la explicación del «desamparo» en seres humanos, se denominó «Reformulación Atribucional de la Teoría del Desamparo Aprendido» (RATDA) (Abramson, L.Y., Seligman, M.E.P., y Teasdale, J., 1978).

Para ambos, como ya anunciamos, es consustancial contemplar en su punto de partida el fenómeno de la *incontrolabilidad* consistente de acuerdo con L.B. Alloy y M.E.P. Seligman (1979), en la experimentación por parte de un organismo de determinados resultados que ocurren independientemente de sus respuestas instrumentales.

*Esta experiencia con «incontrolabilidad» produce en quién la sufre, de acuerdo con lo más básico de la teoría, una expectativa de que no existirá «contingencia» entre su conducta y los resultados en el futuro.* Siendo los efectos de esta expectativa cuatro déficits fundamentales: un motivacional, otro cognitivo, un tercero emocional, y un último de autoestima.

### El Déficit Cognitivo

Básicamente consiste en que un hombre que ha tenido experiencias con «incontrolabilidad», posteriormente tiene dificultades para aprender que sus respuestas, en caso de que las de tendrían éxito, es decir, controlarían la situación.

Las variables que identifican este déficit son dos: El Locus de Control, que habla de expectativa de control y es anterior a la acción; y el Estilo Atribucional, que habla de responsabilidad de la acción y es posterior a ella.

El término *Locus de Control* designa un concepto proveniente de la Teoría del Aprendizaje Social de J.B. Rotter (1954, revisada en 1982). Afirma J.B. Rotter (1966, p. 1) que «cuando un sujeto percibe el refuerzo como resultado de alguna acción propia, pero no siendo enteramente contingente con su acción, entonces, en nuestra cultura, es percibido como el resultado de la suerte, la causalidad, o que está bajo el control de otros poderosos o imprevistos, debido a la cantidad de fuerzas que lo rodean. Cuando el individuo interpreta el suceso de esta forma, designamos esta creencia como una situación de control externo. Si la persona percibe que el suceso es contingente con sus propia conducta o con sus características permanentes, llamamos a esta percepción una creencia en el control interno».

Pues bien, de esto, fundamentalmente, la conclusión que a nosotros nos interesa es la que sostiene una relación de semejanza entre la tendencia cognitiva

llamada «Locus de Control Externo», que implica el creer (expectativa) en la ineeficacia del control personal sobre el refuerzo, y la conducta de los individuos sometidos a situaciones de «desamparo».

El término *Estilo Atribucional* es el rasgo característico más importante de la Reformulación y representa la tendencia de una persona a favor de ciertas categorías causales, al intentar dar sentido a sus experiencias o acontecimientos (Layden, M.A., 1982).

Su origen podemos hallarlo en la Teoría Atribucional de la Motivación de Logro de B. Weiner (1974, 1976), por cuanto en su Reformulación L.Y. Abramson et al. (1978) acudieron a ella para tratar de resolver las dificultades con las que la teoría inicial se encontraba a la hora de explicar el «desamparo» en los seres humanos. L.Y. Abramson et al. (1978) pensaron, como advierte J. Salgado (1984, p. 57), «que con la introducción de las explicaciones causales que las personas realizan cuando se encuentran en situaciones de éxito o fracaso tales dificultades se podrían superar».

En efecto, L.Y. Abramson et al. (1978), primeramente, toman de la Teoría de B. Weiner la idea de que las causas que la gente emplea para explicar sus éxitos y fracasos son las cuatro que B. Weiner propone (capacidad, esfuerzo, suerte y dificultad de la tarea), y que éstas se pueden clasificar en dos dimensiones: «internalidad» y «externalidad», ortogonales entre si y con una tercera que ellos añaden: la «globalidad». Para, seguidamente, sostener que los «desamparados» tienden a realizar de modo sistemático, atribuciones internas, estables y globales para los acontecimientos negativos para sus vidas e, implicitamente, que frente a hechos positivos tienden a manifestar atribuciones externas, inestables y específicas. Es decir, que el «sujeto desamparado», aunque no sea responsable, se atribuye la responsabilidad de sus fracasos, y sus existos los explica como debidos al azar o a la suerte.

## El Déficit Afectivo

El déficit afectivo se concreta en un estado emocional de gran ansiedad. Y, si bien cabe objetar que un aumento de la ansiedad no indica necesariamente que los sujetos presenten problemas emocionales, sin embargo M.E.P. Seligman (1975) sostiene que «en algunos casos» ésta deriva en depresión; es decir, que los «desamparados» en situación de incontrolabilidad presentarán los mismos comportamientos que los depresivos.

De ser ello cierto, tendríamos un cuadro conductual caracterizado por: la pasividad (lentitud o ausencia de respuestas, con manifestaciones como la pérdida de interés, dificultad de concentración, incapacidad de tomar decisiones); las expectativas negativas (fuerte tendencia a evaluar sus propias respuestas como condenadas el fracaso, con expresiones como «es inútil hacer nada»); y los sentimientos de desesperanza e impotencia (creencia de que no son capaces de influir en los aspectos de la vida que tienen alguna importancia para ellos, acompaña-

dos de sentimientos de tristeza y hasta de culpabilidad) (Cfr.: Sobral, J., 1983; Vid.: Beck, A.T., 1967).

### **El Déficit de Autoestima**

El déficit de autoestima surgió porque los autores de la Reformulación adi-  
virtieron que el «desamparo» comenzaba a parecerse a un estado depresivo, por  
lo que si la RATDA aspiraba a convertirse en un modelo explicativo de la de-  
presión se hacía necesario contemplar una de las características que recogen casi  
todas las teorías sobre la misma: la pérdida de autoestima.

Pero, fundamentalmente, es la introducción de la atribución causal, y espe-  
cialmente su dimensión de internalidad, la que propicia la aparición de este  
cuarto déficit del «desamparo» por cuanto es lógico suponer que las atribuciones  
internas del fracaso afecten al propia autoestima, cosa que no sucedería con las  
atribuciones externas.

Básicamente, este déficit propugna que será negativo el juicio que emita de  
sí mismo un sujeto que, tras experimentar ante un tema o situación para él im-  
portante un resultado no contingente con su respuesta, atribuya la responsabili-  
dad de este resultado insatisfactorio a sí mismo, y desarrolle la expectativa de  
que aún cuando los demás si puedan controlar la situación él será incapaz de  
lograrlo.

### **El Déficit Motivacional**

Por último, el déficit motivacional (muy parecido a la inhibición de los de-  
presivos) consiste en la dificultad de los sujetos «desamparados» para emprender  
nuevas conductas, nuevas respuestas voluntarias que, por ejemplo, como el  
escrutinio de su memoria o el cálculo mental den como resultado un buen ren-  
dimiento escolar. Por el contrario, la respuesta típica del niño será «¿para qué»,  
o «¿a quién le importa», o «¿qué más da?» (Bronfenbrenner, V., 1970).

El déficit emocional basicamente imposibilita al sujeto que ha experimenta-  
do la incontrolabilidad para reevaluar y replanificar su conducta, hasta el punto  
de que el «desamparado» carece no sólo de motivación para triunfar sino inclu-  
so, lo que es mucho más grave, para huir del fracaso (Atkinson, J.W., 1964). El  
déficit emocional niega, en definitiva, para el sujeto la viabilidad de cualquier  
proyecto de acción personal.

## **RECURSOS TÉCNICOS**

Uno de los requisitos indispensables a la hora de exponer una investigación  
es el de dejar suficientemente claro los instrumentos utilizados. En este caso nos

limitaremos a mencionar el título de los cuestionarios que hemos aplicado a una muestra de 812 sujetos de octavo de E.G.B.

Para medir el Locus de Control, recurrimos al cuestionario IAR de V.C. Crandall, W. Katlosky, y V.J. Crandall (1965); Para medir el Estilo Atribucional, recurrimos a la adaptación para su aplicación a una población infantil realizada por D. García-Villamisar (1983) del ASQ de C. Peterson et al. (1982); Para medir la Ansiedad, recurrimos al cuestionario IDARE adaptado por TEA a partir de la baremación en España por J. Bermudez Moreno (1977) del original STAI de C.D. Spielberger (1970); Para medir la Autoestima, recurrimos a la escala de M. Rosenberg (1965); Y, por último, para medir la Motivación de Logro, recurrimos al test de J. Smith (1973).

## ALTERNATIVAS DE ANÁLISIS DE LOS DATOS

Básicamente el objetivo de un análisis de datos es poner a prueba las relaciones predichas por la teoría; en nuestro caso, considerando cierto número de variables contextuales (sexo, habitat, clase social, etc.), entre variables de personalidad, variables cognitivas y el fracaso escolar, establecido éste en atención a las calificaciones escolares.

Pues bien, en nuestra opinión, un modelo de análisis del fracaso escolar como el que presentamos, precisa para su contrastación:

a) Revisar si las correlaciones existentes entre las variables mencionados y el fracaso escolar, y del conjunto de las variables entre si, se corresponden con las supuestas por la Teoría. Por ejemplo, a nivel de hipótesis sostiene que el fracaso escolar correlaciona de un modo negativo con la motivación de logro, y que la autoestima correlaciona de un modo negativo con el estilo atribucional interno, estable y global para el fracaso.

b) Analizar si existen diferencias significativas entre los alumnos fracasados y los exitosos en las diferentes variables incluidas en este estudio. Por ejemplo, a nivel de hipótesis se plantea que fracasados y exitosos difieren significativamente en su motivación de logro, evidenciando los segundos un mayor motivación de logro que los primeros.

c) Explorar cuáles son los mejores predictores del fracaso escolar entre nuestras variables y, para ello toda vez que, en nuestro estudio, los grupos (fracasados/exitosos) ya han sido definidos a priori de modo que lo que tendremos que ver es lo específico de cada uno de ellos para poder asignar los individuos a los grupos respectivos, nos inclinamos por la utilización de un análisis discriminante (Martínez Ramos, E., 1984).

## REFERENCIAS BIBLIOGRÁFICAS

- ABRAMSON, L.Y., SELIGMAN, M.E.P., y TEASDALE, J. (1978): «Learned Helplessness in Humans: Critique and Reformulation». *J. Abn. Psychol.*, 87, 49-74.
- ALLOY, L.B., y SELIGMAN, M.E.P. (1979): «On the cognitive component of the learned helplessness». En BOWER, G.H.: *The Psychology of Learning and Motivation*, Vol. 13, Academic Press, New York.
- ATKINSON, J.W. (1964): *An introduction to motivation*. D. Van Nostrand, New York.
- BECK; A.T. (1967): *Depression: clinical, experimental and theoretical aspects*. Hoeber, New York.
- BERMUDEZ MORENO, J. (1977): *Ansiedad y Rendimiento*. Tesis Doctoral. Madrid.
- BRONFENBRENNER, U. (1970): «Psychological costs of quality and equality in education». En ALLAN, V.L. (Ed.): *Psychological factors in poverty*. Markham, Chicago.
- CRANDALL, V.C., KATLOSKY, W., y CRANDALL, V.J. (1965): «Children's beliefs in their control of reinforcements in intellectual academic achievement behaviors». *Chil. Develop.*, 33, 643-661.
- GARCIA VILLAMISAR, D. (1983): *Análisis experimental de un modelo cognitivo procesual interactivo de desamparo aprendido en adolescentes*. Tesis Doctoral. Universidad de Santiago.
- LAYDEN, M.A. (1982): «Attributional style therapy». En ANTAKI, C., y BREWIN, C. (Eds): *Attributions and the psychological change*. Academic Press, Londres.
- MARTÍNEZ RAMOS, E. (1984): «Fundamentos del análisis discriminante y su aplicación en un estudio electoral». En SANCHEZ CARRION, J.J. (Coor.): *Introducción a las técnicas de análisis multivariables aplicadas a las ciencias sociales*. Centro de Investigaciones sociológicas, Madrid.
- MILLER, I.W., y NORMAN, W. (1979): «Learned helplessness in humans: A review and attribution-theory model». *Psycho. Bull.*, 86, 1, 93-118.
- PETERSON, C., SEMMEL, A., von Baeyer, C., ABRAMSON, L.Y., METALSKY, G.I., y SELIGMAN, M.E.P. (1982): «The attributional style questionnaire». *Cog. Ther. Res.*, 6, 3, 287-300.
- RODRÍGUEZ ESPINAR, S. (1982): *Factores de rendimiento escolar*. Oikos-tau, Barcelona.
- ROSEMBERG, M. (1965): *Society and the adolescent self-image*. University Press, Princeton.
- ROTTER, J.B. (1966): «Generalized expectancies for internal versus external control of reinforcement». *Psychol. Monographs*, 80, 1, (todo el n.º 609).
- ROTTER, J.B. (1982): «Social Learning Theory». En FEATHER, N.T. (Ed.): *Expectations and Actions: Expectancy-value Models in Psychology*. Erlbaum, Hillsdale.
- SALGADO, J. (1984): *Desamparo Aprendido: Atribución causal y juicios de similitud*. Tesis Doctoral. Universidad de Santiago.
- SELIGMAN, M.E.P. (1975): *Helplessness*. Freeman, San Francisco. (Existe traducción al castellano: *Indefensión*. Ed. Debate, Madrid, 1981).

- SELIGMAN, M.E.P., ABRAMSON, L.Y., SEMMEL, A., y von BAEYER, C. (1979): «Depressive attributional style». *J. Abn. Psychol.*, 88, 3, 242-247.
- SMITH, J. (1973): A quick measure of achievement motivation». *Br. J. Soc. Clin. Psychol.*, 12, 137-143.
- SOBRAL, J. (1983) *Desamparo aprendido: Una perspectiva desde el Locus de Control y la Motivación de Logro*. Tesis Doctoral. Universidad de Santiago.
- SPIELBERGER, C.D. (1970): *STAI. Manual for State-trait Anxiety Inventory (Self-Evaluation Questionnaire)*. Consulting Psychologist Press, Palo Alto, California.
- WEINER, B. (Ed.) (1974): *Achievement Motivation and attribution theory*. General Learning Press, Morristown.
- WEINER, B., NIEREMBERG, R., y GOLDSTEIN, M. (1976): «Social Learning (Locus of control) versus attributional (Causal stability) interpretations of expectancy of success». *J. Pers.*, 44, 52-68.
- WORTMAN, C.B., y BREHM; J. (1975): «Response to uncontrollable outcomes: An integration of reactance theory and the learned helplessness model». En BERKOWITZ, L. (Ed.): *Advances in experimental Social Psychology*, Vol. 8, Academic Press, New York.

## Revisión del constructo ambiente como variable de proceso

por  
*Luis Miguel Villar*

### EL AMBIENTE DE APRENDIZAJE EN EL AULA COMO LÍNEA DE INVESTIGACIÓN

El estudio del ambiente es una de las áreas de investigación educativa que mayor importancia están cobrando a nivel internacional.

Las influencias ambientales que más se han investigado han sido las familiares, sociológicas e instruccionales y los efectos que ordinariamente se buscan son medidas de rendimiento. Las variables ambientales se han cruzado, además, con otras –generalmente psicológicas– para medir las interacciones entre ellas y su efecto combinado en el rendimiento.

Se debe considerar que hay *dos fuentes* que inciden en al investigación actual sobre el clima de clase: de una parte, la contribución teórica que autores como Murray, Lewin y, posteriormente, Pace y Stern han ejercido para comprender el ambiente y explicar la conducta del individuo. De otra, Chávez, quien ha reivindicado que la evaluación de las percepciones se apoya en el trabajo de otros autores e incluyen métodos observacionales de baja inferencia con los que se mide el clima de clase. (Fraser, B.J. 1985).

### INSTRUMENTOS DE MEDIDA DEL AMBIENTE DE CLASE EN LOS NIVELES NO UNIVERSITARIOS

#### 1. «Inventario de Ambiente de Aprendizaje»

El «Inventario de Ambiente de Aprendizaje», (en inglés LEI), contiene 15 escalas o conceptos predictores del rendimiento, que son relevantes en la teoría e investigación psicosocial, se han hallado útiles en la teoría e investigación educativas o que han sido juzgados intuitivamente como relevantes para la psicología social de la clase. Cada escala contiene siete declaraciones o items.

A través del LEI se puede:

- a) evaluar las percepciones de un estudiante individual, y/o
- b) calcular el ambiente de aprendizaje de la clase como grupo.

Se utiliza la *media* de clase cuando se quiere conocer las variaciones en el enseñanza y en el currículum, y el *sujeto* cuando a uno le preocupa el sexo, el autoconcepto o la personalidad del estudiante.

Estos 15 constructos y su significación experimental son los siguientes:

1. *Cohesividad*. Las clases pequeñas son más cohesivas que las grandes. Las clases enseñadas por profesores inexpertos fueron más cohesivas que aquellas que tenían un profesor familiarizado con el curso. Las clases de historia e inglés son más cohesivas que las de ciencias. La cohesividad está positiva y constantemente relacionada con el aprendizaje.
2. *Diversidad*. No hay nexos entre diversidad y resultados de los estudiantes.
3. *Formalidad*. Las clases grandes son más formales que las pequeñas, y las clases de matemáticas son menos formales que las de ciencias, humanidades y lenguaje. Las relaciones entre formalidad y aprendizaje tienden a ser positivas.
4. *Velocidad*. La velocidad se correlaciona negativamente con la puntuación media de la inteligencia en la clase. Las clases de ciencias y matemáticas son más rápidas que las de inglés e historia. La velocidad no se relaciona con el aprendizaje.
5. *Ambiente material*. Influye enormemente en la estructura del grupo. En un curso de física experimental las evaluaciones del material eran altas cuando había abundancia de materiales y estas puntuaciones se correlacionaban positivamente con varias medidas del aprendizaje del alumno.
6. *Fricción*. La fricción es superior cuando las clases tienen un número superior de niños que de niñas, y se correlaciona negativamente con medidas de aprendizaje.
7. *Dirección de la meta*. Es superior en clases que siguen cursos tradicionales que en cursos usando un programa experimental. Hay una dirección positiva entre la dirección de la meta y el aprendizaje.
8. *Favoritismo*. No se han hallado asociaciones favoritismo-aprendizaje.
9. *Dificultad*. Las clases más grandes se perciben como menos difíciles que las pequeñas. Hay relaciones positivas entre dificultad y aprendizaje.
10. *Apatía*. Esta dimensión complementa la escala de cohesividad. Es consistente y negativamente relacionada con el aprendizaje.
11. *Democracia*. Esta escala no ha discriminado entre cursos, pero se relaciona significativa y positivamente con el aprendizaje.
12. *Pandilla*. Las pandillas son superiores en las clases de cursos tradicionales que en los que siguen un curso experimental, y se asocian negativamente con el aprendizaje.

13. *Satisfacción.* Se relaciona negativamente con el tamaño de la clase y positivamente con el aprendizaje.
14. *Desorganización.* Las clases de matemáticas son percibidas como más desorganizadas que otras, y una desorganización grande está asociada con una reducción en el aprendizaje.
15. *Competitividad.* Es alta en clase con alta proporción de alumnos sobre alumnas, si bien no se han hallado relaciones entre competitividad y resultados.

Por tanto, las escalas que tienen correlaciones positivas y consistentes son las siete siguientes: cohesividad, formalidad, ambiente material, dirección de la meta, dificultad, democracia y satisfacción. (Fraser, B. Anderson, G.y Walberg, H., 1982).

En el Cuadro N.<sup>o</sup> 1 aparecen relacionados estudios realizados con el LEI, en donde se especifican los autores, los resultados esperados, muestra y metología.

## 2. «Inventario de mi clase»

Es un instrumento derivado del LEI, adecuado para niños de 8 a 12 años de edad. Reduce las 15 escalas originales del LEI a sólo cinco (cohesividad, fricción, satisfacción, dificultad y competitividad), y modifica el formato de las hojas de respuesta. Contiene 38 ítems, y con él se ha investigado su validez predictiva en un estudio sobre el rendimiento en la lectura, utilizando a tal fin un análisis de regresión múltiple, tomando la clase como unidad de análisis. Parece que las percepciones de una alta competitividad en la clase se asociaron con puntuaciones inferiores del rendimiento. En otro estudio, Boulanger, F.D. (1980) halló que las percepciones ambientales se asociaban significativamente con resultados de rodamiento. Finalmente, Ellet, C.D., Masters, J.A. y Pool, J.E. (1978) administraron el «Inventario de mi clase» para conocer las relaciones entre el ambiente, rendimiento y asistencia a las clases. Por medio de un análisis de regresión múltiple hallaron que las escalas del MCI explicaron 72 por ciento de la varianza del rendimiento, y 14 por ciento de la varianza del rendimiento de la asistencia a clase.

## 3. «Escala de Ambiente de Clase»

La escala consta de 90 ítems agrupados en nueve subescalas, y estas a su vez, en tres dimensiones: relaciones interpersonales, mantenimiento y cambio de sistema, y orientación a la tarea. Las nueve subescalas miden los siguientes aspectos: Implicación, Afiliación, Ayuda del profesor, Orientación a la tarea, Competición, Orden y Organización, Claridad de normas, Control del profesor e Innovación. (Moos, R.H y David, Th. G. 1981).

## CUADRO N.º 1

Estudio	Resultados de aprendizaje	Muestra				Metodología		
		Localización	Área	Nivel	Tamaño	Unidad de análisis	Variables controladas	Técnicas de análisis
Walberg y Anderson (1968b); Anderson y Walberg (1968); Walberg (1969b; c; 1972)	Rendimiento	EE.UU.; en algunos casos, en Canadá	Física	BUP	Máximo de 144 aulas	Estudiante o clase	Pretest. y C.I.	Correlación simple, múltiple o canónica
O'Reilly (1975)	Rendimiento en matemáticas	Canadá	Matemáticas	9 y 10	48 aulas	Clase	C.I.	Correlación simple, parcial o múltiple
Cort (1979)	Medidas de rendimiento y actitud	11 Estados norteamericanos	Estudios Sociales	5 y 6	81 aulas	Clase	Pretest. tamaño de clase	Correlación múltiple y canónica
Power y Tisher (1979)	Rendimiento	Australia	Ciencias	BUP	20 aulas	Estudios o clase	Pretest	Correlación simple, múltiple o canónica
Walberg, Singh y Rasber (1977)	Rendimiento	India	Ciencias, Ciencias Sociales	10	150 aulas	Clase subgrupo	C.I.	Correlación simple y múltiple
Paige (1973, 1979)	Rendimiento cognitivo	Indonesia	—	6	60 colegios	Clase	Ambiente de casa, «background» del niño	Regresión múltiple

### 3.1. Relación entre la «Escala de Ambiente de Clase» y el rendimiento.

En el estudio de Villar, L.M. (1984) se propusieron entre otras, la hipótesis que declaramos a continuación:

«Existe una relación lineal entre el promedio del rendimiento académico global y las nueve subescalas de la «Escala de Ambiente de Clase».

El procedimiento de cálculo fue el paquete BMDP2R, cuyos resultados sintetizamos en la Tabla N.<sup>o</sup> 1. En esa hipótesis, propusimos los valores medios obtenidos por cada sujeto en cada una de las nueve subescalas. Entendimos que estos valores describían mejor las distintas dimensiones ambientales que pueden tener un peso en el rendimiento. En algunas investigaciones se ha visto cómo el rendimiento se asocia más a determinadas subescalas, y esa razón fue la que nos indujo a conocer el peso de las distintas subescalas en el rendimiento.

Advertimos cómo la ecuación resultante está formada por las subescalas 3 («Ayuda del profesor»), 2 («Afiliación»), 7 («Claridad de normas») y 4 («Orientación a la tarea»), que en su conjunto explican un 50 por ciento de la varianza del rendimiento ( $R^2$ ), teniendo, asimismo, como coeficiente de correlación múltiple (R) el valor .7137. En definitiva, la magnitud de la contribución de estas cuatro subescalas como predictoras del rendimiento puede considerarse relativamente importante.

TABLA N.<sup>o</sup> 1

Variable	Coeficiente	Correlación parcial r	Coeficiente de regresión parcial típico $\beta$
Subescala 3 (``Ayuda del profesor``)	.05683	.5798	.346
Subescala 2 (``Afiliación``)	.04159	.4923	.252
Subescala 7 (``Claridad de normas``)	.04256	.5056	.242
Subescala 1 (``Orientación en la tarea``)	.04947	.2843	.201
Costante	-.50248	Coeficiente de correlación múltiple $R = .7137$	Error estándard de estimación .2238
$Y = .05683 \times \text{Subescala } 3 + .04159 \times \text{Subescala } 2 + .04256 \times \text{Subescala } 7 + .04947 \times \text{Subescala } 4 - (.50248)$			

## EL AMBIENTE DE APRENDIZAJE EN LA ENSEÑANZA UNIVERSITARIA

### 1. «Cuestionario de Ambiente de Clases Individualizadas»

Este instrumento se puede aplicar a alumnos de primer curso de carrera de Facultades y EE.UU. y, en particular, a estudiantes de EE.UU. de Formación del Profesorado de EGB.

Este sistema se distingue de otros instrumentos de ambiente en los siguientes aspectos:

- a) Introduce cinco dimensiones para distinguir una clase individualizada de otra convencional.
- b) Utiliza hojas y cuestionarios de evaluación para las enseñanzas real y preferida o ideal.
- c) Se puede utilizar con profesores y con alumnos.
- d) Está diseñado para que se pueda corregir a mano.
- e) Tiene un formato reducido (50 ítems) para ofrecer una medida económica del ambiente de clase. (Fraser, B.J. 1982).

#### 1.1. Asociaciones entre resultados del estudiante y ambiente de clase

Cuando se ha usado en los niveles no universitarios existe correlación entre el ICEQ y actitudes de los alumnos. En otro aspecto, se halló que las conclusiones y las generalizaciones se asociaban más a la dimensión «investigación», y el interés por el ocio en ciencias se relacionaba con una «participación» superior.

Finalmente, hay relaciones globales entre las percepciones de los estudiantes sobre la individualización y los resultados de final de curso. Las asociaciones son positivas entre el nivel de individualización y el rendimiento en metas actitudinales.

## TIPOS DE INVESTIGACIONES PROPUESTOS CON INSTRUMENTOS DE AMBIENTE

### 1. Asociaciones entre resultados del estudiante y ambiente de clase

Este tipo de investigación controla estadísticamente ciertas características de los estudiantes, especialmente el pretest y la habilidad general. En este tipo de estudios se han usado correlaciones simples, múltiples y canónicas.

Parece un resultado consistente que orden y organización tienen un influencia positiva en el rendimiento en una variedad de metas.

## 2. Las percepciones del ambiente como variables de criterio

- a) Las dimensiones parecen ser adecuados indicadores para la evaluación del curriculum.
- b) Por medio de las dimensiones se pueden analizar las diferencias entre las percepciones de los estudiantes y profesores sobre los ambientes real y preferido o ideal, y.
- c) a través de los instrumentos ambientales se pueden realizar estudios que impliquen otras variables independientes, como personalidad del profesor, tamaño de clase, nivel, materia, ambiente de centro, tipo de escuela, etc.

### REFERENCIAS BIBLIOGRAFICAS

- BOULANGER, F.D. (1980): «Relationship of an inservice program to student learning: Naturalistic documentation», *Science Education*, 64, págs. 349-355.
- ELLETT, C.D., MASTERS, J.A. and POOL, J.E. (1978): «The incremental validity of teacher and student perceptions of school environment characteristics», Paper presented to the Annual Meeting of the Georgia Educational Research Association, Atlanta, January.
- FRASER, B.J. (1985). *Two decades of research on perceptions of classroom environment*, Paper presented at AERA, Chicago.
- FRASER, B.J., ANDERSON, G.J. and WALBERG, H.J. (1982).: *Assessment of learning environments: Manual for Learning Environment Inventory (LEI) and My Class Inventory (MCI)*, Western Australian Institute of Technology, Bentley.
- FRASER, B.J. (1982).: *ICEQ. Individualized classroom environment questionnaire. Handbook*, Australian Council for Educational Research.
- MOOS, R.H. and DAVID, Th. G. (1981).: «Evaluating and Changing Classroom Settings», en EPS-TEIN, J. (Ed.): *The Quality of School Life*, D.C. Heath and Company, Lexington, Massachusetts, págs. 59-80.
- VILLAR ANGULO, L.M. (Dir.) (1984). *Calidad de enseñanza y supervisión instruccional*, ICE de la Universidad, Sevilla.

## RELACIÓN DE AUTORES Y CENTROS DE TRABAJO

Apellidos y Nombre	Centro de Trabajo	
Anguera Argilaga, M. Teresa	Dpto. Psicología Experimental	Barcelona
Arnal Agustín, Justo	Escuela Univ. Profes. EGB	Barcelona
Bartolomé Pina, Antonio R.	Dpto. Pedagogía Experimental	Barcelona
Bartolomé Pina, Margarita	Dpto. Pedagogía Experimental	Barcelona
Benito Diez-Canseco, José	Escuela Univ. Prof EGB	Oviedo
Bisquerra Alcina, Rafael.	Dpto. Pedagogía Experimental	Barcelona
Bordas Alsina, Inmaculada	Dpto. Metodología y T. Educativa	Barcelona
Buendía Eisman, Leonor	Sección de Pedagogía	Granada
Cabrera Rodríguez, Flor	Dpto. Pedagogía Experimental	Barcelona
Cajide Val, José	Sección de Pedagogía	Santiago
Colás Bravo, M.ª Pilar	Sección de Pedagogía	Sevilla
Corral Blanco, Norberto	Sección de Pedagogía	Oviedo
Diez Miralles, Manuel	Facultad de Medicina	Alicante
Echeverría Samanes, Benito	Dpto. Pedagogía Experimental	Barcelona
García Jiménez, Eduardo	Sección de Pedagogía	Sevilla
García Llamas, J. Luis	Dpto. Pedagogía Experimental	Madrid-Uned
Gómez Benito, Juana	Dpto. Psicología Experimental	Barcelona
Jiménez Fernández, Carmen	Dpto. Pedagogía Experimental	Madrid-Uned
Lizasoain Hernández, Luis	Sección de Pedagogía	S. Sebastián
López Blasco, Andrés	Diputación	Valencia
Loscos Lucero, M.ª Paz	Dpto. Pedagogía Experimental	Madrid-Uned
Marín Gracia, M.ª Ángeles	Dpto. Pedagogía Experimental	Barcelona
Marqués Ruiz, Jesús	Diputación	Valencia
Martínez González, Raquel	Sección de Pedagogía	Oviedo
Martínez Sánchez, Amparo	Dpto. Didáctica	Valencia
Mateo Andrés, Joan	Dpto. Pedagogía Experimental	Barcelona
De Miguel Díaz, Mario	I.C.E. Universidad	Oviedo
De la Orden, Arturo	Dpto. Pedagogía Experimental	Madrid
De Pablos Pons, Juan	Sección de Pedagogía	Sevilla
Pastor Ciurana, Jesús T.	Facultad de Ciencias	Alicante
Peris Pascual, M.ª Dolores	Dpto. Pedagogía Experimental	Valencia
De Pro Bueno, Antonio	Escuela Univ. Prof. EGB	Melilla
Rajadell Puigros, Nuria	Dpto. Pedagogía Experimental	Barcelona
Rodríguez Espinar, Sebastián	Dpto. Pedagogía Experimental	Barcelona
Rodríguez Lajo, Mercedes	Dpto. Pedagogía Experimental	Barcelona
Rosales López, Carlos	Sección de Pedagogía	Santiago
Rubal Rodríguez, Xosé	Sección de Pedagogía	Santiago
Salvador Mata, Francisco	Escuela Univ. Prof. EGB	Granada
Sánchez Algarra, Pedro	Dpto. Pedagogía Experimental	Barcelona
Seisdedos, Nicolás	Ediciones T.E.A.	Madrid
Tejedor Tejedor, F. Javier	Sección de Pedagogía	Santiago
Trillo Alonso, Felipe	Sección de Pedagogía	Santiago
Villar Angulo, L. Miguel	Sección de Pedagogía	Sevilla

## DIRECCIONES CENTROS DE TRABAJO

<i>Facultad de Ciencias</i>	C/S. Vicente de Raspeig	03004-ALICANTE
<i>D. Metod. y Tecn. Educativa</i>	FACULTAD FIL. CC. EDUCACIÓN Baldiri Reixach s/n. Bq. «D», 4. <sup>o</sup>	08028-BARCELONA
<i>D. Pedag. Experimental, Terapéutica y Orientación</i>	Baldiri Reixach s/n. Bq. «D», 3. <sup>o</sup>	08028-BARCELONA
<i>D. Psicología Experimental</i>	Baldiri REixach s/n. Bq. «D», 2. <sup>o</sup>	08028-BARCELONA
<i>Escuela Univ. Prof. EGB</i>	C/. Melchor de Palau, 140	08014-BARCELONA
<i>Sección de Pedagogía</i>	FACULTAD FIL. CC. EDUCACIÓN C/. Campus de Cartuja	18011-GRANADA
<i>Escuela Univ. Prof. EGB</i>	C/. Gran Vía, 56	18010-GRANADA
<i>D. Pedag. Experimental y Orientación</i>	FACULTAD FIL. CC. EDUCACIÓN Edif. «B». Ciudad Universitaria	28003-MADRID
<i>D. Pedag. Experimental</i>	UNIV. NACIONAL DISTANCIA Ciudad Universitaria	28040-MADRID
<i>Ediciones T.E.A.</i>	C/. Fray Bernardino de Sahagun, 24	28036-MADRID
<i>Escuela Univ. Prof. EGB</i>	Avda. Alfonso XIII s/n.	MELILLA
<i>Sección de Pedagogía</i>	FACULTAD FIL. CC. EDUCACIÓN C/. Valdes Sala s/n.	33007-OVIEDO
<i>I.C.E.</i>	C/. Quintana 30, 1. <sup>o</sup>	33009-OVIEDO
<i>Escuela Univ. Prof. EGB</i>	C/. Aniceto Sela, s/n.	33005-OVIEDO
<i>Sección de Pedagogía</i>	FACULTAD FIL. CC. EDUCACIÓN C/. Alto de Zorroaga, s/n.	20014-S. SEBASTIÁN
<i>Sección de Pedagogía</i>	FACULTAD FIL. CC. EDUCACIÓN Avda. Juan XXIII, s/n.	SANTIAGO
<i>D. Pedagog. Experimental</i>	FACULTAD FIL. CC. EDUCACIÓN C/. Gonzalo de Bilbao, 7-9	41003-SEVILLA
	FACULTAD FIL. CC. EDUCACIÓN Avda. Blasco Ibáñez, 21	46010-VALENCIA

# REVISTA INVESTIGACION EDUCATIVA

## BOLETIN DE DOMICILIACION BANCARIA.

## BOLETIN DE SUSCRIPCION

**Para suscribirse llene este boletín y devuélvalo a:**  
**REVISTA DE INVESTIGACIÓN EDUCATIVA**  
Dpto. de Pedagogía Experimental, Terapéutica y Orientación  
Facultad de Ciencias de la Educación  
Baldíri Reixach s/n. Bloque D - piso 3.<sup>o</sup>  
08028-BARCELONA

Nombre .....

Institución .....

Dirección .....

Población ..... C.P. ....

País ..... Teléfono .....

Marque con una cruz los números que desea recibir:

1983: N.<sup>o</sup> 1  N.<sup>o</sup> 2 (Extra)  1.000 ptas.

1984: N.<sup>o</sup> 3  N.<sup>o</sup> 4 (Agotado) 500 ptas.

1985: N.<sup>o</sup> 5  N.<sup>o</sup> 6 (Extra)  1.000 ptas.

1986: N.<sup>o</sup> 7  N.<sup>o</sup> 8  1.500 ptas.

(Fecha y Firma)

Señores,

Les agradeceré que con cargo a mi cuenta/libreta atiendan los recibos que les presentará P.P.U. Promoción Publicaciones Universitarias, como pago de mi suscripción a la REVISTA DE INVESTIGACIÓN EDUCATIVA

Titular de la cuenta .....

Banco/Caja .....

N.<sup>o</sup> de cuenta ..... N.<sup>o</sup> de libreta .....

Agencia .....

Población .....

(Fecha y Firma)

1