

Sumario

Editorial.....	3
La investigación sobre evaluación educativa..... <i>Leonor Buendía Eisman</i>	5
La evaluación de los alumnos en el contexto de la evaluación de la calidad de las universidades..... <i>Fuensanta Hernández Pina</i>	25
La evaluación en la educación secundaria obligatoria en el nuevo sistema educativo en Andalucía..... <i>Eugenio Hidalgo Díez</i>	51
La evaluación del profesorado. Un tema a debate..... <i>Joan Mateo, Tomás Escudero, Mario de Miguel, José Ginés Mora y Sebastián Rodríguez</i>	73
Las ayudas del profesor en el aprendizaje cooperativo y su influencia en la evaluación criterial..... <i>Leonor Buendía Eisman</i>	95
Evaluación de la resolución de problemas aritméticos en Primaria..... <i>Enrique Castro, Luis Rico, José Gutiérrez, Encarnación Castro, Isidoro Segovia, Nicolás Morcillo, Francisco Fernández, Evaristo González y Antonio Tortosa</i>	121
Pruebas estandarizadas y evaluación del rendimiento: usos y características métricas..... <i>J. M. Jornet Meliá y J. M. Suárez Rodríguez</i>	141



Revista de Investigación Educativa

Volumen 14, número 2, 1996

Revista de Investigación Educativa

Volumen 14, número 2, 1996

Asociación Interuniversitaria de Investigación Pedagógica (AIDIPE)

EVALUACIÓN EDUCATIVA

Leonor Buendía Eisman (Coordinadora)



Revista de Investigación Educativa

Volumen 14, número 2, 1996

Asociación Interuniversitaria de Investigación Pedagógica (AIDIPE)

Directora:

Fuensanta Hernández Pina

Directora ejecutiva:

Julia Victoria Espín

CONSEJO ASESOR:

Margarita Bartolomé

Nuria Borrell

Leonor Buendía

Iñaki Dendaluce

Lisardo Doval

Narciso García

Fuensanta Hernández Pina

Jesús Jornet

Mario de Miguel

Arturo de la Orden

Ramón Pérez Juste

Antonio Rodríguez Diéguez

Francisco J. Tejedor

Carmen Vidal

CONSEJO DE REDACCIÓN:

Manuel Álvarez

Rafael Bisquerra

Carmen Buisan

Flor Cabrera

Inmaculada Dorio

Julia Victoria Espín

Pilar Figuera

Javier Gil

M^a Ángeles Marín

Fuensanta Hernández

M^a Luisa Rodríguez

Mercedes Rodríguez

Delio del Rincón

M^a Paz Sandín

SUSCRIPCIÓN E INTERCAMBIO CIENTÍFICO:

Dpto. de Métodos de Investigación y Diagnóstico en Educación

Facultad de Pedagogía

Paseo del Valle de Hebrón, 171, 2^a planta

08035 BARCELONA (Spain)

DISTRIBUCIÓN:

Dpto. de Métodos de Investigación y Diagnóstico en Educación

Facultad de Educación

Campus de Espinardo

Universidad de Murcia 30100

Tel. (968) 36 40 67

DISEÑO Y MAQUETACIÓN:

Compobell, S.L. MURCIA

ISSN: 0212-4068

Depósito Legal: MU-724-1996

Revista de Investigación Educativa

Volumen 14, número 2, 1996

Editorial	3
La investigación sobre evaluación educativa	5
<i>Leonor Buendía Eisman</i>	
La evaluación de los alumnos en el contexto de la evaluación de la calidad de las universidades	25
<i>Fuensanta Hernández Pina</i>	
La evaluación en la educación secundaria obligatoria en el nuevo sistema educativo en Andalucía	51
<i>Eugenio Hidalgo Díez</i>	
La evaluación del profesorado. Un tema a debate	73
<i>Joan Mateo, Tomás Escudero, Mario de Miguel, José Ginés Mora y Sebastián Rodríguez</i>	
Las ayudas del profesor en el aprendizaje cooperativo y su influencia en la evaluación criterial	95
<i>Leonor Buendía Eisman</i>	
Evaluación de la resolución de problemas aritméticos en Primaria	121
<i>Enrique Castro, Luis Rico, José Gutiérrez, Encarnación Castro, Isidoro Segovia, Nicolás Morcillo, Francisco Fernández, Evaristo González y Antonio Tortosa</i>	
Pruebas estandarizadas y evaluación del rendimiento: usos y características métricas	141
<i>J. M. Jornet Meliá y J. M. Suárez Rodríguez</i>	



E D I T O R I A L

En la reunión de la Junta Directiva de AIDIPE, celebrada en Barcelona, en febrero de 1996, se acordó editar un nuevo número monográfico de la Revista de Investigación Educativa. Se realizaron bastantes propuestas y se conjugaron diferentes criterios. Se optó finalmente por realizarlo sobre Evaluación Educativa, tanto por la trayectoria que tiene la Asociación en investigaciones sobre el tema, como por la oportunidad e importancia en el momento actual.

Un monográfico sobre evaluación, obliga a seleccionar unos aspectos y olvidar otros, y en último punto perder el sentido global e interactivo de los diferentes elementos que intervienen en el sistema educativo. ¿Cómo centramos en la evaluación de los alumnos/as, o del aula, sin evaluar los proyectos curriculares, el contexto del Centro, el equipo directivo, el profesor/ra etc.? No obstante, la selección era necesaria; así pues, procuramos recoger aspectos fundamentales que puedan marcar líneas de trabajo para seguir profundizando en el tema y en segundo lugar investigaciones empíricas sobre evaluación en un área concreta.

En el primer artículo se realiza una revisión de los trabajos sobre evaluación de un importante número de revistas. Posteriormente se incluyen dos trabajos teóricos; uno sobre los alumnos universitarios en el contexto de la evaluación de la calidad de las Universidades y el otro sobre Educación Secundaria Obligatoria.

La evaluación del profesorado, junto con una propuesta de construcción de pruebas de evaluación, completan el bloque de trabajos teóricos realizados por especialistas en cada uno de los temas.

En la segunda parte de la revista se presentan dos estudios empíricos. El primero sobre evaluación en Educación Primaria, a partir de pruebas criteriales, y el segundo sobre la evaluación de las dificultades de los alumnos/as en la resolución de problemas aritméticos.

Esperamos que estas aportaciones, nos ayuden a conocer un poco mejor la situación de la investigación sobre evaluación educativa y abran líneas de trabajo para poder seguir profundizando en el tema.

LA INVESTIGACIÓN SOBRE EVALUACIÓN EDUCATIVA

Leonor Buendía Eisman
Universidad de Granada

RESUMEN

En este estudio se realiza una revisión sobre los trabajos de evaluación y posteriormente se reflexiona sobre tres aspectos:

- a) Evaluación formativa versus sumativa*
- b) Papel de la Evaluación en la Enseñanza*
- c) El conocimiento de los profesores/las sobre evaluación.*

ABSTRACT

In this study is accomplished a review on the evaluation projects and afterwards is reflected on three aspects:

- a) Formative Evaluation versus Sumative*
- b) Paper of the Evaluation in the Teaching*
- c) The Knowledge of the Teachers on Evaluation.*

I. INVESTIGACIÓN SOBRE EVALUACIÓN EDUCATIVA

La evaluación ha sido y sigue siendo un tema preferente en educación desde diferentes perspectivas y con diferentes finalidades. A veces es una cuestión política,

Leonor Buendía Eisman. Catedrática de Universidad. Área MIDE, F. de Ciencias de la Educación. Campus de la Cartuja. Universidad de Granada. lbuedia@platon.ugr.es

otras social y la mayoría de las veces, para los profesores/as e investigadores, un continuo problema cuando se plantean explicar, comprender o mejorar la educación.

En este trabajo, hemos comenzado por revisar los artículos de una selección de revistas que suelen incluir investigaciones sobre el tema, y una vez revisadas incidir en aquellos aspectos que resultan de mayor interés y a la vez de mayor repercusión en la práctica educativa.

Las revistas seleccionadas han sido:

- A.E. Acción Educativa
- A.E.H.E. Assessment & Evaluation in Higher Education
- A.I.E. Aula de Innovación Educativa
- B. Bordón
- B.E.R.J. Britis Educational Research Journal
- C.P. Cuadernos de Pedagogía
- E.T.R.D. Educational Technology Research and Development
- ELT. J. ELT. Journal
- E.J.P.E. European Journal of Psychology of Education
- H.E. Higher Education
- R.I.E. Revista de Investigación Educativa
- R.I.F.P. Revista Interuniversitaria de Formación del Profesor/aado
- S.E.E. Studies in Educational Evaluation

AUTOR (año) Y REVISTA	TEMA	TIPO ESTUDIO	MUESTRA
Abraira, C.F. (1995) B.	Evaluación de dos programas educativos, utilizando el modelo CIPP.	Estudio cuasi-experimental con pretest y postest.	Dos grupos de alumnos/as de magisterio. N= 91 y 78 Universidad de León.
Abraira, C.F. (1995) R.I.E.	Aplicación del modelo CIPP con alumnos/as de magisterio para mejorar el rendimiento en Matemáticas.	Estudio empírico. Técnicas observacionales (escalas de estimación) y entrevistas.	522 alumnos/as de Magisterio de la Universidad de León.
Al-karni, A. (1995) H.E.	Evaluación de Dptos. universitarios.	Cuestionario.	46 Dptos. de 7 universidades de Arabia Saudí.
Alvik, T. (1995) S.E.E.	Evaluación de la clase.	Investigación-acción. Estudio de casos.	Centros de Oslo.
Argos, J. y Ezquerra, M.P. (1995) R.I.F.P.	Evaluación del Proyecto Curricular.	Teórico.	
Ballester, J. (1995) C.P.	Evaluación del trabajo en grupo. Autoevaluación.	Teórico/práctico.	
Batallos, J.M. (1995) A.I.E.	Evaluación democrática.	Teórico.	

AUTOR (año) Y REVISTA	TEMA	TIPO ESTUDIO	MUESTRA
Bax, S. (1995) <i>ELT.J.</i>	Evaluación del profesorado.	Estudio de casos: entrenamiento de profesores/as.	Profesores/as Sudafricanos.
Bennett, S.J., and others. (1995) <i>A.E.H.E.</i>	Evaluación del rendimiento lector en Química.	Grupo de discusión. Auto-evaluación con cuestionarios.	Alumnos/as de 1º y 2º año.
Bradford, I.W. (1995) <i>A.E.H.E.</i>	Taxonomías de Evaluación del Aprendizaje.	Teórico. Análisis de contenido de autores.	
Buendía, L.; Fernández, A. y Ruiz, J. (1995) <i>R.I.E.</i>	El estudio plantea tres líneas de trabajo. La primera sobre evaluación de la práctica docente, la segunda sobre evaluación de las nuevas tecnologías en educación y la tercera sobre evaluación de la investigación.	Estudios teóricos con propuestas de aplicación a la práctica.	
Contreras, E. (1995) <i>B.</i>	Derechos de los alumnos/as.	Teórico.	Alumnos/as universitarios
de Miguel, M. (1995) <i>A.E.</i>	Evaluación de programas. Influencia de la participación ciudadana en la política educativa.	Teórico.	
Escudero, J.M. (1995) <i>A.I.E.</i>	Evaluación del Proyecto Curricular.	Teórico.	
Fernández, J. (1995) <i>C.P.</i>	Funciones de la evaluación. Reflexión sobre el concepto.	Teórico.	
Fullada, J. y Pallisera, M. (1995) <i>A.I.E.</i>	Evaluación de procedimientos: habilidades básicas y habilidades perceptivo-motrices.	Pruebas criterios.	Niños/as 4 años escolarizados. Adultos con disminución psíquica.
García, J.M. (1995) <i>B.</i>	Evaluación institucional.	Teórico.	Universidad privada.
Gosling, P. (1995) <i>E.J.P.E.</i>	Rendimiento global según opinión de los profesores/as.	Cuestionarios.	Profesores/as secundaria.
Gregoire, J. and Nieuwen, C. (1995) <i>E.J.P.E.</i>	Evaluación de contenidos: matemáticas.	Cuasi-experimental.	Niños/as de 5 y 6 años.
Gutiérrez, J. (1995) <i>R.I.E.</i>	Revisión de trabajos de evaluación en educación ambiental.	Meta-análisis cualitativo sobre trabajos de educación ambiental.	Actas de congresos. Tesis doctorales. Revistas internacionales.
Marjoribanks, K. (1995) <i>S.E.E.</i>	Implicación de los padres en la evaluación.	Correlacional	516 familias australianas.
McChesney, P. (1995) <i>B.J.G.C.</i>	Evaluación de programas de orientación para la carrera.	Cualitativo-etnográfico.	Estudiantes universitarios de ciencias y geografía.

AUTOR (año) Y REVISTA	TEMA	TIPO ESTUDIO	MUESTRA
Norton, L. and Crowley, C. (1995) H.E.	Evaluación de programas de aprendizaje.	Análisis de contenido. Concepciones sobre aprendizaje.	Alumnos/as de 1º de psicología.
O'Neill, T. (1995) S.E.E.	Evaluación de proyectos.	Teórico.	
Psacharopoulos, G. (1995) I.R.E.	Evaluación del rendimiento global entre países.	Cuasi-experimental. Pruebas estandarizadas.	Centros públicos y privados: EEUU, Canadá, Corea, España e Irlanda.
Ramírez, L. y Rodríguez, A. (1995) C.P.	Evaluación de contenidos: Lenguaje.	Entrevistas y cuestionarios.	
Rodríguez, B. (1995) B.	Evaluación del proyecto curricular de centros (modelo evaluación interna).	Teórico. Autoevaluación.	
Ruggles, C. (and others) (1995) I.E.T.I.	Evaluación de recursos (programas de ordenador para enseñanza)	Cuestionarios Investigación descriptiva.	Estudiantes universitarios de Geografía y Educación.
Rul, J. (1995) A.I.E.	Concepto de evaluación.	Teórico.	
Sainsbury, M. and others (1995) B.E.R.J.	Evaluación de estándares.	Teórico.	
Scriven, M. (1995) S.E.E.	Evaluación del profesor/aado (costos de la evaluación personal).	Teórico.	
Serrano, J.; Tojar, J.C.; Galante, R. y Elósegui, E. (1995) R.I.E.	Tiene dos partes. La primera sobre metodología en la construcción de instrumentos y sistemas de indicadores. En la segunda parte se realiza una revisión de los trabajos del 1994 sobre toma de decisiones y política educativa.		Teórico
Stuart, M. (1995) B.J.E.P.	Evaluación de contenidos: Lengua.	Longitudinal.	Niños/as de 5 ó 6 años.
Stufflebeam, D. (1995) S.E.E.	Evaluación de profesores/as y directivos.	Descriptivo-comparativo. (modelos de evaluación personal).	
Webster, W. (1995) S.E.E.	Modelo para la evaluación de programas, de escuelas y alumnos/as.	Pretest-postest. Cuasi-experimental.	120 centros.
Weston, C. (and others). (1995) E.T.R.D.	Evaluación formativa. Concepto evaluación.	Empírico: análisis de contenido. (textos).	11 textos sobre evaluación formativa de los últimos 20 años.
Bastianoni, P.; Scappini, E. and Emilianí, F. (1996) E.J.P.E.	Estudia los procesos de cambio en los niños y la relación entre una conducta inadaptada y la práctica de los adultos (según un programa previamente diseñado).	Metodología observacional	24 observaciones de cada niño por cada maestro/a y 24 auto-observaciones de cada adulto sobre sus actos con cada niño.

AUTOR (año) Y REVISTA	TEMA	TIPO ESTUDIO	MUESTRA
Kwan, K-P. and Leung, K. (1996) A.E.H.E.	Evaluación por dos métodos diferentes, (el tutor y el grupo de iguales) del rendimiento de los alumnos/as en sesiones de simulación como profesores/as.	Estudio comparativo a partir de una escala de puntuación.	5 grupos, cada uno con 17-20 estudiantes.
Lewy, A. (1996) S.E.E.	Describe tres categorías en la Evaluación del rendimiento y la situación actual, diferenciando Evaluation y Assessment	Estudio teórico y sugerencias para construir test de aula (tareas de evaluación).	Escuelas primarias de Israel.
Meyer, T. y Anis, J. (1996) E.J.P.E.	Evalúan la calidad de las respuestas humanas a cuestiones sobre un texto, comparándolas con las respuestas artificiales de un programa de ordenador.	Cuasi-experimental. 2 grupos (comparación en diferentes variables de respuestas humanas vs. máquinas).	332 sujetos. (cuestiones sobre significado de un texto y parafrasear el texto.
Pacek, D. (1996) E.L.T. Journal.	Evaluación de un programa de Enseñanza del Inglés.	Evaluación posttest a través de cuestionarios. Estudio teórico.	Escuelas japonesas.
Shaw, K. E. (1996) A.E.H.E.	Influencia de los aspectos culturales en la Evaluación.	Estudio teórico.	
Sumsion, J. and Fleet, A. (1996) A.E.H.E.	Evaluación de profesores/as en formación sobre el desarrollo de la práctica reflexiva.	Escala de observación: altamente reflexivo, mode- radamente reflexivo, no reflexivo.	124 estudiantes a profesores/as de los cuales sólo 73 perma- necieron en el estudio com- pleto.
Toohy, S.; Rian, G. y Hyghes, C. (1996) A.E.H.E.	Se identifican diferentes modelos de evaluación del prácticum y se hacen propuestas para el mejor aprovechamiento del mismo.	Estudio teórico con impor- tantes aportaciones del prácticum al desarrollo profesional.	
Willian, D. (1996) B.E.R.J.	Tiene dos partes: a) Revisión teórica de la validez de constructo; b) Aplicaciones a la evaluación del currículum nacional. Aplicación del modelo a los problemas que se presentan.	Estudio teórico.	
Willian, D. and Black, P. (1996) B.E.R.J.	Estudio sobre Evaluación Formativa y Sumativa, funciones de cada una y efectos en el aprendizaje.	Estudio teórico.	
Works, S. and Barak, M. (1996) S.E.E.	Evaluación interna y externa de un programa interdisciplinar de ciencia y tecnología.	Cuestionario y grupos de discusión y visitas de supervisores.	En el norte de Israel, 17 escuelas.
Zabala, A. (1996) A.I.E.	Estudio sobre los informes de evaluación, el contenido de los mismos y la utilidad de la información.	Estudio teórico.	Informes a padres, pro- fesores/as, alumnos/as y administración.

AUTOR (año) Y REVISTA	TEMA	TIPO ESTUDIO	MUESTRA
Boijaard, D. and Verloop, N. (1996) S.E.E.	Describe las características del conocimiento práctico de los profesores/as y problemas asociados con la evaluación de este conocimiento.	Estudio teórico. Repercusiones sobre los sistemas de evaluación del profesor/a.	
Berberoglu, G.; Dochy, F. and Moerkerke, G. (1996) E.J.P.E	Importancia de evaluar el conocimiento previo para conseguir un buen aprendizaje.	Estudio comparativo entre hombres/mujeres en diferentes áreas: matemáticas, ciencias sociales y ciencias.	Educación superior. Universidad de Turquía.
Chartier, D. (1996). E.J.P.E.	Evaluación de una serie de dimensiones cognitivas en dos grupos de adultos que siguen dos programas de aprendizaje diferente.	Empírico. Metodología comparativa. (diseño cuasi-experimental con pretest y postest).	250 adultos, 80% hombres.
Cousins, B.J. (1996) S.E.E.	Evaluación participativa. Los efectos de la investigación-acción en la evaluación y práctica educativa.	Estudios de casos con observación participante.	3 Escuelas públicas de Canadá. (1995).
Dochy, F.; Moerwerke, G. and Martens, R. (1996) S.E.E.	Estudio detallado del conocimiento previo, presentando un modelo de integración de Aprendizaje, Instrucción y Evaluación.	Análisis de los efectos del conocimiento previo sobre rendimiento y otras variables. Revisión teórica.	Se analizan 129 estudios.
Ellett, Ch. D. and others. (1996) S.E.E.	Evaluación de profesores/as a través de los estándares de Evaluación personal de Stufflebeam (1988).	Empírico (Escala) (STAR).	45.000 Profesores/as de Escuelas públicas.
Galante, R y Sanchidrian, C. (1996) R.I.E.	Evaluación de los contenidos de los trabajos de tesis y tesinas en Historia de la Educación y una comparación entre estos y los artículos de la Revista de Hª de la Educación.	Estudio bibliométrico.	Años: 1980-1990, Universidades, n = 16; Tesis, n = 169; tesinas, n = 219.
Hager, P. and Butler, J. (1996) A.E.H.E.	Contrasta el modelo de evaluación basado en medición objetiva y la evaluación desde la Teoría constructivista.	Estudio teórico con diferentes análisis comparativos de ambos modelos.	
Hargreaves, D.J.; Galton, M.J. and Robinson, S. (1996) E.R.	Elaboración de una taxonomía de los constructos o vocabulario de evaluación utilizados por los profesores/as de primaria para evaluar a alumnos/as en arte, música y creación literaria.	Observaciones y entrevistas utilizando a profesores/as. Estudio empírico.	25 descripciones de actividades para evaluar el arte. 24 descripciones en música. 49 descripciones en creación literaria.
Kniveton, B.H. (1996) A.E.H.E.	Opiniones de los estudiantes sobre evaluación continua y exámenes.	Estudio comparativo por edad y género a través de cuestionario.	292 alumnos/as, estudiantes de Ciencias Sociales de dos Universidades.

En esta revisión los aspectos seleccionados, según los criterios anteriormente comentados, han sido:

- a) Evaluación sumativa versus formativa
- b) Papel de la evaluación en el proceso de enseñanza
- c) Momentos de la Evaluación
- d) El conocimiento que tienen los profesores/as sobre evaluación

2. EVALUACIÓN DE RESULTADOS VERSUS EVALUACIÓN FORMATIVA

Los modelos de evaluación están estrechamente vinculados a las diferentes elaboraciones teóricas sobre las situaciones educativas y el papel que el profesor/a y alumnos/as desempeñan en cada una de ellas. El profesor/a a la hora de diseñar situaciones educativas que favorezcan el aprendizaje, no es lo mismo apoyarse en una teoría que concibe el aprendizaje como un proceso constructivo que en concepciones innatas o en el modelo de procesamiento de la información.

La perspectiva teórica del profesor/a y sus concepciones sobre la educación en general, orientarán en gran medida el modelo de evaluación que el profesor/a realice en sus clases. La evaluación será una valoración de resultados finales con el fin de tomar decisiones institucionales, sociales o políticas, o un elemento más del proceso de enseñanza, encaminado al conocimiento del alumnos/a para la mejora educativa.

En los últimos años los trabajos realizados sobre evaluación del proceso de enseñanza-aprendizaje han proliferado hasta el punto que en la revisión mostrada por Sechrest y Figueredo (1993) sobre las revistas *Evaluation Review* y *New Directions in Program Evaluation*, encontraron que en cinco números del volumen 15 de la ER, sólo dos artículos tenían como primer objetivo el informe de los resultados de una evaluación, quince fueron fundamentalmente metodológicos y uno estaba referido a las condiciones para utilizar los resultados de una evaluación. En la *New Directions in Program Evaluation* no suelen publicar resultados de evaluaciones.

El interés por las evaluaciones centradas en el proceso educativo, de carácter cualitativo, ha aumentado considerablemente. El número de miembros de la *American Evaluation Association* es el mismo, según Sechrest (1992), en el grupo de evaluación cuantitativa que en el grupo de evaluación cualitativa.

En esta perspectiva de evaluación, se encuentran entre otros, el modelo de evaluación respondiente de Stake (1983), la evaluación iluminativa de Parlett y Hamilton (1983), la evaluación democrática de McDonald (1983) y los modelos centrados en la evaluación del cambio (Schwille y otros, 1986) etc.

En todos ellos se pretende conocer por qué se ha producido un determinado resultado. La finalidad es explicativa, encaminada a la mejora del desarrollo del proyecto educativo.

La evaluación de resultados cumple una finalidad estrictamente demostrativa (Kidder y Judd, 1986); estima la magnitud de los efectos del programa educativo sobre el colectivo que se aplica, respondiendo con ello a la conjetura realizada antes de su aplicación sobre los previsibles efectos. Dentro de esta perspectiva objetivista (Wor-

then y Sanders, 1987) podemos encontrar modelos diferentes en cuanto a los referentes de evaluación (evaluación basada en objetivos, Tyler 1950, el modelo de Cronbach, 1982; el modelo de toma de decisiones de Stufflebeam, 1971), pero todos requieren que la información sea científicamente objetiva (Farley, 1985). Esta se consigue utilizando instrumentos válidos y fiables; y los datos, pueden ser reproducidos y verificados por otros profesionales, acabando el proceso en el momento que los datos son obtenidos.

Desde esta perspectiva, la evaluación del alumno/a ha sido utilizada como justificación de un programa educativo (Cronbach y Snow, 1969; Scriven, 1967; Stake, 1967; Stufflebeam, 1969), sin considerar específicamente al profesor/a y en muchos de los casos deliberadamente olvidado como parte del proceso, por la consideración tecnicista del papel del evaluador (Webster, 1995).

Con los resultados, además de analizar el impacto de un programa educativo, se pretenden conocer si se han cubierto los objetivos de política educativa, relacionada con estándares de rendimiento, recayendo sobre el resultado académico del alumno/a gran parte de la responsabilidad de todo el funcionamiento de un sistema.

En estos casos el problema se plantea en el establecimiento de criterios para comparar resultados finales entre países, como el comentado por Linn y Baker (1995), recogiendo los resultados ofrecidos por la IAEP sobre el rendimiento en Matemáticas de cinco países: Korea, Hungría, Taiwan, Italia y Estados Unidos. Los criterios seguidos fueron los presentados en el informe de 1994 de la NCEST (National Council on Education Standards and Testing) para evaluar el sistema educativo:

- a) Contenido estándar (definido como aquello que debería enseñarse y que los estudiantes deberían aprender)
- b) Rendimiento estándar (definido como el nivel mínimo que los estudiantes deberían alcanzar)
- c) Oportunidades de aprendizaje (entendido como la adecuación y posibilidades de oportunidades de aprendizaje que se ofrecen a los alumnos/as).

Con estos criterios se realizaron comparaciones entre grupos de iguales y entre países, estableciendo cotas y analizando la situación porcentual del rendimiento medio de cada país respecto a esa cota.

3. PAPEL DE LA EVALUACIÓN EN LA ENSEÑANZA

En el aula tienen lugar una serie de actuaciones tanto individuales, por parte de profesores/as y alumnos/as, como compartidas por ambos, encaminadas a conseguir el desarrollo de las capacidades previstas en los objetivos, la construcción del conocimiento y la formación de actitudes, valores y normas. En la escuela, el alumno/a, puede no sólo rendir académicamente sino además aprender las normas de conducta socialmente más apropiadas.

Aunque el aprendizaje individual es el objetivo primero para la escolarización, sin embargo la educación formal es esencialmente un proceso social (Weinstein, 1991). En este proceso social, las diferencias individuales y culturales influyen en el desarrollo

de las competencias de los alumnos/as. Indudablemente las dificultades aumentan cuando hay discontinuidad entre la cultura de la escuela y de la casa (Cazden, 1988). La necesidad de partir de estas diferencias, de no considerar la escuela como un espacio homogéneo para aprender (todos parten del mismo punto y elaboran de la misma manera, tienen los mismos significados, etc.) exige plantear la evaluación desde una perspectiva diferente al establecimiento de estándares para la evaluación de un producto.

En este apartado nosotros nos vamos a referir al aprendizaje profundo (Marton, 1984; Entwistle and Entwistle, 1991) o aprendizaje con entendimiento (Harlen and James, 1996), implicando esta noción una contextualización del conocimiento en estrecha relación con personas, cosas, ideas o hechos que han tenido lugar en la vida real, y cómo evaluarlo.

El rápido aumento de conocimientos, junto con la inestabilidad de las situaciones de trabajo, ha hecho que pierdan su importancia determinados aprendizajes de hechos y procedimientos, aplicables solamente a situaciones cerradas del aula, para hablar de escuelas de contenido universal (Broudy, 1988).

Crooks (1988) plantea dos formas diferentes de enfrentarse con el aprendizaje según se pretenda un aprendizaje real (Harlem y James, 1996), profundo, o, un aprendizaje rutinario, superficial. En el aprendizaje profundo hay una intencionalidad por comprender lo que se aprende y una conexión entre el conocimiento nuevo con conocimientos previos y experiencia; en el superficial, la intencionalidad se encamina a reproducir el contenido como se requiere, aceptando pasivamente las ideas e informaciones. Todo el aprendizaje se focaliza sobre los requerimientos de la evaluación para la promoción.

Los profesores/as que tienen una visión de escuela «universal», en palabras de Broudy, enseñarán a los alumnos/as los procedimientos adecuados para aprender y la posibilidad de generalizar lo aprendido de un ámbito a otro.

Para dar respuesta a ambas demandas, Marton (1984) propone una categoría intermedia: «aprendizaje estratégico», haciendo referencia con esta categoría a las situaciones escolares en las cuales un aprendizaje eficaz es a menudo una combinación de aprendizaje superficial y aprendizaje profundo. Si pretendemos que el alumno/a aprenda cada cosa en profundidad, posiblemente aprendería tan pocas cosas que la propia sociedad demandaría más contenidos; por el contrario, un aprendizaje superficial no tendría posibilidad de ser aplicado y el valor de dichos aprendizajes sería nulo.

Un aprendizaje estratégico permite al alumno/a captar las exigencias de la tarea y responderla adecuadamente. En esta situación los procedimientos seguidos para realizar la tarea, son contenidos que hay que aprender y saber aplicar. Esto exige al profesor/a, enseñar al alumno/a qué debe aprender y cómo debe aprenderlo para dar respuesta satisfactorias a las demandas escolares y a la vez conectarlo con experiencias previas o situaciones de la vida real. No es tanto enseñarle a hacer algo por hacer (finalidad del acto en sí mismo) sino como un medio para enseñarles a aprender.

Ayudar al alumno/a a utilizar la cantidad de información facilitada por los diferentes medios que influyen sobre el niño, se ha convertido en un objetivo de primer orden. El alumno/a, con la ayuda del profesor/a, le da sentido a los nuevos conoci-

mientos que construye, gracias a los significados de los conocimientos que ya tiene. La creación de estas conexiones en la escuela dependen tanto de la participación activa del aprendiz, como de la mediación que realiza el profesor/a (Buendía, 1993a), y de la familiaridad con el contexto (Harlem y James, 1996) y los materiales (Weston y McAlpine, 1994).

La evaluación tiene un papel distinto en cada tipo de aprendizaje.

En la evaluación de aprendizajes superficiales como fechas, hechos memorísticos en general o habilidades físicas, la evaluación se utiliza sólo para conocer si dichos conceptos han sido aprendidos o no, y el feed-back proporciona información de lo no memorizado o erróneamente recordado.

Cuando se aprende algo con comprensión, el aprendiz lo entiende e interioriza activamente. Este aprendizaje, vinculado a una experiencia previa, puede ser aplicado a situaciones diferentes de las que enmarcaron su aprendizaje. Lo que es conocido y aprendido podría, por supuesto, cambiar con nuevas experiencias y nuevas ideas, estableciéndose conexiones para darle sentido y significatividad a lo nuevo.

La evaluación de este aprendizaje permitiría conocer el punto de partida del aprendiz y el proceso seguido en la elaboración o construcción de los nuevos aprendizajes. La evaluación no es aquí una herramienta de control sino una parte importante del proceso de enseñanza.

4. MOMENTOS DE LA EVALUACIÓN EN EL PROCESO DE ENSEÑANZA

Sin duda una de las mayores discusiones sobre el tema se han realizado sobre los modelos de evaluación, intentando establecer relaciones funcionales entre el tipo de evaluación que se realiza y la opción ideológica personal; y si bien esto tiene parte de verdad (no creemos que ningún acto humano este desprovisto de valores y mucho menos cuando implica toma de decisiones), el problema se plantea cuando lo que subyace en la discusión es un desconocimiento técnico y de las diferentes bondades de las opciones en función de los fines a los que sirven (Wilian y Black, 1996).

Decimos esto porque la evaluación final v.s. continua v.s. inicial, puede ser objeto de discusión cuando la opción, por una de ellas (sólo evaluación final para evaluar resultados) responde a una concepción de la evaluación como justificación social del nivel alcanzado por el alumno/a y demostración de los actos «objetivos» del profesor/a. Pero, si por el contrario, en función del momento en el que se realiza, optamos por una u otra, el problema es más una cuestión temporal que propiamente de adscripción teórica o ideológica.

La evaluación inicial tiene por objeto realizar un diagnóstico para detectar el grado de desarrollo en aspectos básicos del aprendizaje y será el punto de referencia del tutor o tutora para la toma de decisiones relativas al desarrollo de currículo. Consiste, por tanto, en recoger información acerca de las situación actual de cada alumno/a al iniciar un determinado proceso de enseñanza-aprendizaje con objeto de adecuar este proceso a sus posibilidades.

Fases de enseñanza aprendizaje objeto de evaluación Inic.	Contenido de la evaluación	Momento
Comienzo de la escolaridad	Recabar datos sobre el desarrollo respecto a su etapa evolutiva. Datos médicos, psicopedagógicos o familiares. Datos obtenidos por observación directa del alumno/a.	Al comienzo de la escolaridad.
Comienzo de cada momento de aprendizaje	Análisis del punto de partida de cada alumno/a de las capacidades para el aprendizaje. Conocimiento de las ideas previas.	Antes de desarrollar la planificación.
Comienzo de una unidad didáctica	Activar las ideas previas. Comprobar la adecuación de los nuevos aprendizajes a las características del alumno/a y al contexto del Centro.	Antes de concretar la unidad didáctica a desarrollar.
Comienzo de un ciclo	Estudio de los datos de escolarización anterior. Estudio y análisis de los Informes de Evaluación Individualizados. Análisis del punto de partida para nuevos aprendizajes.	Al iniciar el ciclo.
Comienzo de una etapa	Estudio de los datos de escolarización anterior. Estudio y análisis de los Informes de Evaluación individualizados. Grado de desarrollo de los aprendizajes en los distintos ámbitos.	Al iniciar la etapa.
Al escolarizar a un niño procedente de un traslado	Estudio y análisis de: a) informe de evaluación individual extraordinario. b) Informe de evaluación de años anteriores de escolarización Recabar información del grado de desarrollo de los aprendizajes en las diferentes áreas.	Al ingresar en el Centro.

(1) Elaborado a partir de documentos del CEP de Granada.

La evaluación continua tiene como finalidad detectar y analizar los progresos y dificultades de los alumnos/as, indagar en sus causas y en consecuencia realizar la intervención educativa.

Es formativa en tanto que es un elemento más del proceso educativo que proporciona la información gracias a la cual dicho proceso puede retroalimentarse y corregirse constantemente. Los objetivos didácticos constituyen el punto de referencia inmediato de la evaluación continua.

Su finalidad es:

Conocer el grado de desarrollo de las capacidades expresadas en los objetivos generales de etapa y/o de ciclo.

Conocer el grado de desarrollo de los objetivos, contenidos y criterios de evaluación de las distintas áreas del currículo en cada contexto.

Adaptación de los elementos curriculares a las características del alumno/a y al contexto del Centro.

Analizar las dificultades encontradas para desarrollar los objetivos y la asimilación de los contenidos.

Conocer el aprendizaje realizado por el alumno/a y el criterio que del mismo tiene (autoevaluación).

Conocer el esfuerzo, atención, interés y motivación en cada secuencia de aprendizaje.

En esta fase evaluativa el profesor/a investiga la marcha del proceso de enseñanza-aprendizaje. Es a través de las evaluaciones de los aprendizajes de los alumnos/as como los profesores/as indagan, se hacen preguntas sobre su papel en el proceso seguido. A partir de aquí, la metodología utilizada, las estrategias de enseñanza, las tareas desarrolladas y todas las variables referidas al alumno/a, deben ser reconsideradas en un proceso de indagación sistemática que permita siguiendo las palabras de Cochran-Smith y Susan Lytle **aprender de la enseñanza**.

«Cada clase es vista como una situación de investigación y una fuente de conocimiento más accesible cuando los profesores/as reflexionan o se preguntan conjuntamente sobre las cuestiones surgidas en ellas, enriqueciendo su teoría de la práctica» (1993: 51).

La evaluación final debe considerarse como la consecuencia de la evaluación continua. Tiene por finalidad conocer el grado de desarrollo de capacidades a lo largo de un tramo de enseñanza-aprendizaje (nivel, ciclo, curso, etc.) y las dificultades encontradas para comenzar el siguiente tramo educativo. Debe integrarse como el primer elemento de la evaluación inicial o diagnóstica en un nuevo proceso educativo.

Con frecuencia se confunde el rol de este momento de la evaluación y el papel que el profesor/a debe desempeñar en él, con la evaluación de los resultados de un programa ajeno al centro, e incluso al aula, en el que el profesor/a asumen un rol técnico, como si de un externo al proceso de enseñanza-aprendizaje se tratara (Sechrest and Figueredo, 1993). Sin embargo en la evaluación educativa, la evaluación final es un momento más de la espiral de la evaluación formativa. Al igual que la continua, no

los dilemas prácticos que se les plantean para llevar a término los propósitos educativos en estas situaciones».

Las situaciones de Enseñanza-Aprendizaje son siempre situaciones de indagación-acción. En el aula tienen lugar una serie de interacciones, interactividades (Buendía y Pegalajar, 1990; Buendía, 1993a) y procesos de gran similitud con cualquier proceso investigador; lo que ha permitido ver la enseñanza como una situación constante de indagación, y los procesos, que tienen lugar en la situación educativa, como actos intencionales encaminados a la construcción de un conocimiento útil tanto para el profesor/a (conocimiento práctico: Fenstermacher, 1994; Shön, 1992, Clandinin y Connelly, 1988; Elbaz, 1981), como para el alumno/a (aprendizaje). La figura del profesor/a investigador (Stenhouse, 1975), no es la del profesional que investiga y aplica los conocimientos generados en contextos fuera del aula sino la del investigador en la acción, el que transforma la enseñanza en una continua situación de indagación-acción-reflexión y toma de decisiones. El propio proceso de desarrollo curricular debe ser concebido como una investigación del profesor/aado, «un modo poderoso por el que los profesores/as comprenden cómo ellos y sus alumnos/as construyen y reconstruyen el currículum» (Cochram-Smith y Lytle, 1993: 51).

Las teorías constructivistas, basadas en la idea de que los humanos crean y construyen activamente su realidad personal enfatizan los procesos de conocimiento y se orientan hacia la evaluación de la viabilidad (utilidad) como opuesto a validez (verdad) de la visión del mundo única de cada individuo (Neimeyer 1996).

Sin embargo, pocos profesores/as son conscientes del conocimiento que tienen de las situaciones educativas. Una de las mayores dificultades con las que se encuentran es unir su actuación docente y las decisiones que se toman en el aula, con la terminología que, desde la formalización teórica del conocimiento sobre la situación educativa, se utiliza. Con frecuencia hemos podido constatar que sólo los separa el código de lenguaje utilizado para expresar las prácticas profesionales, siendo esto, a veces, motivo de la gran separación existente entre teóricos y prácticos. Este conocimiento personal, elaborado desde la propia práctica, de hacerse consciente y explícito, permitiría disponer de un importante cuerpo teórico en la formación de los profesores/as y un marco de referencia sobre el conocimiento base de la enseñanza (Buendía, 1995).

5. BIBLIOGRAFÍA

- ABRAIRA, C.F. (1995). ¿La Evaluación Formativa ayuda a los alumnos/as a aprender? *Bordón*, 47, 1, 53-67.
- ABRAIRA, C.F. (1995). Incidencia de la evaluación formativa en el proceso didáctico. *Revista de Investigación Educativa*, 25, 79-90.
- AIRASIAN, P. (1991). Perspectives on measurement instruction for preservice teachers, *Educational Measurement: Issues and Practice*, 10 (1), 13-16.
- AL-KARNI, A. (1995). Evaluating the performance of academia departement chaipersons. *Higher education*. 29, 1, 37.
- ALVIK, T. (1995). School-based evaluation: a close-up. *Studies in educational evaluation*. 21, 3, 311.

es un conjunto de puntuaciones de diferentes resultados de aprendizaje. Concebir así la evaluación continua o la evaluación final supone convertir el aula en una continua situación de examen; y la evaluación educativa en una continua evaluación sumativa.

5. EL CONOCIMIENTO DE LOS PROFESORES/AS SOBRE EVALUACIÓN

Ni el planteamiento de modelos teóricos, concibiendo las situaciones educativas como superposición de variables cada vez de mayor complejidad, ni el conocimiento formal, generado desde instancias fuera de la escuela, han sido suficientes para ayudar a los maestros/as y profesores/as al cambio conceptual que exige una evaluación con finalidades formativas y por lo tanto como una parte más de la enseñanza.

La mayoría de los profesores/as, en los cursos de formación, demandan conceptos e instrumentos que le permitan mejorar sus prácticas docentes y más concretamente la evaluación de los aprendizajes (Gullickson, 1986), por no considerar relevante para su tarea, ni útil para resolver los problemas que se le presentan, la bibliografía que se le recomienda. (Gullickson, 1993; Schafer y Lissitz, 1987; Stiggins y Faires-Conklin, 1988).

Algunos de los conceptos y habilidades para la evaluación, tales como son presentados en los libros de medición (Hanna, 1993; Linn y Gronlund, 1995; Mehrens y Lehmann, 1991; Nitko, 1996; Oosterhof, 1996; Salvia y Ysseldyke, 1995), son derivados de un modelo que comienza en el siglo pasado, amparado en el método científico. La instrucción en la clase es considerada equivalente al tratamiento; y la teoría de los test sugiere que las herramientas de evaluación son diseñadas para evaluar cuidadosamente los resultados de la instrucción en las diferentes materias.

Algunos profesores/as, amparándose en los criterios de validez y fiabilidad desde la perspectiva psicométrica, dicen hacer la evaluación según un modelo que se podría caricaturizar con la perspectiva de Galton (1889), según la cual, los estudiantes difieren en su inherente capacidad de aprender el contenido de las diferentes disciplinas. El evaluador es, en esta perspectiva, el científico que puede objetivamente evaluar y recordar, de cada estudiante, los resultados de su rendimiento. Los estudiantes son el foco de observación, y el modelo de medición presupone que ellos se comportan como objetos pasivos. Sin embargo esto, realmente, no es así. El profesor/a no es ese científico ni actúa de esa manera; y el alumno/a, no es un objeto. El alumno/a participa en todo el proceso, interactúa con el profesor/a y con sus iguales cuando construye todo tipo de contenidos; y al igual que el profesor/a, observa, valora, emite juicios, reflexiona, corrige sus actos etc. El profesor/a en la clase, en contraste científico experimental, es un observador participante (Whyte, 1943) o con palabras de Vidich y Lyman «un investigador etnográfico», y por lo tanto, capaz de generar un conocimiento construido desde las prácticas, contextualizado, y para los mismos fines desde los que fue generado. Otra cosa bien distinta es que el profesor/a sea consciente que sus decisiones son producto de esta construcción y como tal lo explicita.

Sin duda, la necesidad de crear un cuerpo de conocimientos científicos para aplicarlos a la enseñanza, desde el paradigma positivista, se ha desplazado hacia la necesidad de crear un conocimiento del profesor/a y de su práctica y que Carter (1990:299) lo define como «el conocimiento que tienen los profesores/as sobre las situaciones de clase y

- ARGOS, J. y EZQUERRA, M.P. (1995). La evaluación del proyecto curricular: ¿Tarea rutinario-burocrática o proceso de reflexión compartida? *Revista interuniversitaria de Formación del Profesorado*. 22, enero-abril, 139-150.
- BALLESTER, J. (1995). Evaluación del trabajo en Grupo. *Cuadernos de Pedagogía*. 234, marzo, 48-54.
- BASTIANONI, P.; SCAPPINI, E.; EMILIANI, F. (1966). Children in residential care: How to evaluate behavioral change. *European Journal of Psychology of Education*. XI, 4, 459-471.
- BATALLOSO, J.M. (1995). ¿Es posible una evaluación democrática? O sobre la necesidad de evaluar educativamente. *Aula de Innovación Educativa*, Febrero. 35, 73-77.
- BAX, S. (1995). Principles for evaluating teacher development activities. *ELT journal*. 49, 3, 262.
- BEIJAARD, D. y VERLOOP, N. (1966). Assessing Teachers' Practical Knowledge. *Studies in Educational Evaluation*. 22, 3, 275-286.
- BENNETT, S. J. (1995). «HELPA»: a rapid means of student evaluation of lecturing performance in higher education. *Assessment & evaluation in higher education*. 20, 2, 191.
- BERBEROGLU, G.; DOCHY, F.; MOERKERKE, G. (1966). Psychometric evaluation of entry assessment in higher education: A case study. *European Journal of Psychology of education*. XI, 1, 25-42.
- BRADFORD, I.W. (1995). Assessment for learning: quality and taxonomies. *Assessment & evaluation in higher education*. 20, 2, 171.
- BROUDY, H.S. (1988). *The Uses of Schooling*. New York: Routledge.
- BROWN, M. (1996). Formative assessment for learning. En Black, P. and Michel, A. (Eds.): *Assessment in Schools: International Comparisons*. Berkeley, Ca: UCLA.
- BUENDÍA, L. (1993a). Los mecanismos de influencia educativa en el logro de los objetivos de Educación Primaria. Investigación realizada para el segundo ejercicio de cátedra de Universidad. Universidad de Granada.
- BUENDÍA, L. (1993b). Diseño de una evaluación criterial de aula en el área de Lengua y de Conocimiento del Medio. Actas de: *Jornades sobre avaluació a l'aula*. ICE. Universidad de Barcelona. 181-187 y 244-250.
- BUENDÍA, L. y PEGALAJAR, M. (1990). El análisis de secuencias para la evaluación de la interacción. *Revista de Investigación Educativa*, 8, 16, 457-461.
- BUENDÍA, L. y FERNÁNDEZ, A. (1991). Una experiencia exploratoria en cognición numérica en escolares de 3º de E.G.B. *Bordon*. 41 (1).
- BUENDÍA, L.; FERNÁNDEZ, A. y RUIZ, J. (1995): Estudios evaluativos en diferentes contextos. *Revista de Investigación Educativa*. 26, 159-185.
- BUENDÍA, L. (1995). Formación del profesor/ado universitario. Ponencia presentada en el Congreso Nacional sobre Orientación y Evaluación Educativa. Universidad de la Coruña.
- BUENDÍA, L. (1985). Progreso académico en función del profesor/a. *Revista de Investigación Educativa*. 3 (5), 25-33.
- BUENDÍA, L. y SALMERÓN, H. (1994). Construcción de pruebas criterioles de aula. *Revista de Investigación Educativa*. 23, 405-411.

- CARTER, K. (1990): Teachers' knowledge and learning to teach. In Houston, W. R. (Ed.): *Handbook of Research on Teacher Education*. New York: Macmillan; 291-310.
- CARTER, K. (1990): Teacher's knowledge and learning to teach. In W. Houston (ed.) *Handbook on Research of Teacher Education*, pp. 291-310. New York. Macmillan.
- CASTEJÓN, F.J. (1995). Evaluación de la condición física: una propuesta. *Aula de Innovación Educativa*, Junio, 39, 59-64.
- CLANDININ, D.J., & CONNELLY, F.M. (1988). Teacher's personal knowledge: Whats counts as «personal» in studies of the personal. *Journal of Curriculum Studies*. 19, 487-500.
- COCHRAN-SMITH, M. and LYTLE, S.L. (Eds.) (1993): *Inside/Outside: Teacher Research an Knowledge*. New York: Teachers College Press.
- CONTRERAS, E. (1995). Derechos de los alumnos/as universitarios ante la evaluación de sus aprendizajes. *Bordón*. 47, 1, 31-41.
- COUSINS, J.B. (1966). Consequences of Researcher Involvement in Participatory Evaluation. *Studies in Educational Evaluation*. 22, 1, 3-28.
- CRONBACH, L. (1982): *Designing evaluations of educational and social programs*. Jasssey-Bass. San Francisco.
- CRONBACH, L.J. y SNOW, R.E. (1969). *Individual differences in learning ability as a funtion of instruccional variables*. Stanford, California. Stanford University Press.
- CROOKS, T.J. (1988). The impact of classroom evaluation practices on pupils. *Review of Educational Reseach*, 58(4) 438-481.
- CHARTIER, D. (1966). An Evaluation of two cognitive learning methods in adults on pre-qualification schemes: Logo and logical reasoning workshops (ARL). *Europen Journal of Psychology of Education*. XI, 4, 443-457.
- DE MIGUEL, M. (1995). Participación ciudadana en la evaluación de las políticas públicas. *Acción Educativa*. 12, 66. 211-218.
- DOCHY, F.; MOERKERKE, G. and MARTENS, R. (1966). Integrating Assessment, Learning and Instruction: Assesment of Domain-Specific and Domain-Transcending Prior Knowledge and Progress. *Studies in Educational Evaluation*. 22, 4, 341-362.
- ELBAZ, F. (1981): *The Teacher's Practical Knowledge*. Londres: Croom Helm.
- ELLET, C.D. and others. (1966). Looking Backwards with the Personal Evaluation Standards: An Analysis of the Development and Implementation of a Statewide Teacher Assessment Program. *Studies in Educational Evaluation*. 22, 1, 79-90.
- ENTWISTLE, N.J. and ENTWISTLE, A.C. (1991). Forms of understanding for degree examinations: the pupilexperience and its implications. *Higher Education*, 22, 205-227.
- ESCUADERO, J.M. (1995). La evaluación del Proyecto Educativo de Centro. Qué evaluar y cómo hacerlo. *Aula de Innovación Educativa*, Mayo. 38, 41-45.
- FARLEY, J. y otros (1985): *Reconceptualization of vocational education program evaluation*. Universidad de Ohio.
- FENSTERMACHER, G.D. (1994). The Knower and the known: The nature of Knowledge in research of teaching. *Review of Research in Education*. 20, 3-56.
- FERNÁNDEZ, J. (1995). ¿Evaluación? No, gracias, calificación. *Cuadernos de Pedagogía*. 243, Enero, 92-98.

- FULLADA, J. y PALLISERA, M. (1995). Aplicación de las pruebas criterioles en la evaluación de los procedimientos. *Bordón*. 47, 1, 41-53.
- FULLANA, J. (1996). La investigación sobre variables relevantes para la prevención del fracaso escolar. *Revista Investigación Educativa*. 14, número 1. 63-90.
- GALANTE, R. y SANCHIDRIAN, B. (1996). Evaluación de los trabajos de grado en Historia de la Educación desde una perspectiva bibliométrica (1980-1990). *Revista de Investigación Educativa*. v.14, 1, 37-62.
- GARCÍA, J.M. (1995). La Evaluación Institucional a través del Departamento de Recursos Humanos en una Institución universitaria privada. *Bordón*. 47, 1, 17-31.
- GALTON, F. (1889). *Natural inheritance*. New York. MacMillan.
- GOSLING, P. (1995). Description vs exploration and academic evaluation. *European journal of psychology of education*. 10, 1, 41.
- GREGOIRE, J. and NIEUWEN, C. (1995). Counting at nursery school and primary school: toward an instrument for diagnostic assesment. *European journal of psychology of education*. 10, 1, 61.
- GULLICKSON, A.R. (1993). Matching measurement instruction to classroom-based evaluation: Perceived discrepancies, needs, and challenges. En S. L. Wixie y J.C. Conoley (eds.), *Teacher training in measurement and assessment* Measurements, NE: Burros Institute of Mental Measurement, University of Nebraska.
- GULLICKSON, A.R. (1986). Teacher education and teacher percived needs in educational measurement and evaluation. *Journal of Educational Measurement*, 23 (4), 347-354.
- GUTIÉRREZ, J. (1995). Los estudios evaluativos en el ámbito de la educación ambiental. *Revista de Investigación Educativa*. n. 26, 185-200.
- HAGER, P. y BUTLER, J. (1966). Two Models of Educational Assessment. *Assessment & Evaluation in Higher Education*. 21, 4, 367-378.
- HANNA, G.S. (1993). *Better Teaching Through Better Measurement*. Fort Worth, TX: Harcourt Brace Javanovich College Publishers.
- HARGREAVES, D.J.; GALTON, M.J. and ROBINSON, S. (1966). Teachers' assessment of primary children's classroom work in the creative arts. *Educational Reseach*, 38,2, 199-211.
- HARLEN, W. and JAMES, M. (1996) *Assessment and learning: Differences and relationships between formative and summative assessment*. Paper presented in the BERA policy Task Group on Aessment Symposium at European Educational Research Conference, ECER 96. Sevilla.
- KIDDER, L. y JUDD, C.M. (1986). *Research methods in social relations*. Holt, Rinehart and Winston. Nueva York.
- KNIVETON, B.H. (1966). Student perception of assessment methods. *Assessment & Evaluation in Higher Education*. 21, 3, 229-237.
- KWAN, K-P. and LEUNG, R. (1966). Tutor versus peer group assessment of student performance in a simulation training exercise. *Assessment & Evaluation in Higer Education*. 21, 3, 205-213.
- LEWY, A. (1966). Postmodernism in the Field of Achievement Testing. *Studies in Educational Evaluation*. 22, 3, 223-244.

- LINN, R.L. y GRONLUND, N.E. (1995). *Measurement and Assessment in Teaching*. Englewood Cliffs, NJ: Prentice Hall.
- LINN, R.L. y BAKER, E.L. (1995). What do international assessment imply for world-class standards. *Educational Evaluation and Policy Analysis*, 17, 4, 405-418.
- MARJORIBANKS, K. (1995). Parents' involvement in learning as an opportunity structure, a model for evaluation. *Studies in educational evaluation*. 21, 1, 73.
- MARTON, F. and other. (1984). *The Experience of Learning*. Edinburgh: Scottish Academic Press.
- McCHESNEY, P. (1995). Yes, but how?: development and evaluation of careers education in higher education by careers advisers. *British journal of guidance & counselling*. 23, 3, 327.
- McDONALD, B. (1983): La evaluación y el control de la educación. En J. Gimeno y A. Pérez (Eds.): *La enseñanza: su teoría y su práctica*. Akal Universitaria. Madrid.
- McDONALD, B. (1992): *Microworlds and real worlds. An agenda for evaluation*. The European Conference about Information Technology in Education: A Critical Insight. Barcelona. (Documento policopiado).
- MEHRENS, W.A. y LEHMANN, I.J. (1991). *Measurement and Evaluation in Education and Psychology*. Fort Worth, TX: Holt, Rinehart, and Winston, Inc.
- MEYER, T. y ANIS, J. (1966). Assessment of answers to questions about a text and knowledge of cognitive processes. *European Journal of Psychology of Education*. XI, 3, 313-328.
- NEIMEYER, G.J. (1996). *Evaluación constructivista*. Paidós. Barcelona.
- NITKO, A.J. (1996). *Educational Assessment of Students*. Englewood Cliffs, NJ: Prentice Hall.
- NORTON, L.S. and CROWLEY, C. M. (1995). Can students be helped to learn how to learn? An evaluation of an approach to learning programme for first year degree students. *Higher education*. 29, 3, 307.
- OOSTERHOF, A. (1996). *Developing and Using Classroom Assessments*. Englewood Cliffs, NJ: Prentice Hall.
- O'NEILL, T. (1995). Implementation frailties of Guba and Lincoln's fourth generation evaluation theory. *Studies in educational evaluation*. 21, 1, 5.
- PACEK, D. (1966). Lessons to be learnt from negative evaluation. *ELT Journal*. 50, 4, 335-343.
- PARLETT, M. y HAMILTON, D. (1983). La evaluación como iluminación. En J. Gimeno y A. Pérez (Eds.): *La enseñanza: su teoría y su práctica*. Akal Universitaria. Madrid.
- PÉREZ JUSTE, R. (1992). *El modelo evaluativo de la Reforma Educativa*. UNED. Madrid.
- PSACHAROPOULOS, G. (1995). Evaluation of education and training: What room for the comparative approach? *International Review of Education*, 41, 3-4, 259-284.
- RAMÍREZ, L. y RODRÍGUEZ, A. (1995). «INTERNIVELES». Autoevaluación en Lengua. *Cuadernos de Pedagogía*. 241, noviembre, 38-43.
- RODRÍGUEZ, B. (1995). Hacia un modelo de Evaluación interna del Currículo del Centro Educativo. *Bordón*. 47, 1, 5-17.
- ROVINELLI, R.J. y HAMBLETON, R.K. (1977). On the use content specialist in the

- assessment of criterion-referenced test item validity. *Dutch Journal of Educational Research*, 2, 49-60.
- RUGGLES, C.; UNDERWOOD, J. and WALKER, D. (1995). Stile: a hypermedia approach to the provision of a flexible, interdisciplinary resource base for teaching and learning. *Innovations in education and training international*. 32, 3, 209.
- RUL, Jesús (1995). La evaluación comunicativa, factor de desarrollo humano, organizativo y curricular. *Aula de Innovación Educativa*, junio. 39, 70-77.
- SADLER, R. (1989): Formative assessment and design of instructional systems. *Instructional Science*, 18, 119-144.
- SAINSBURY, M. and other. (1995). Issues in the Evaluation of Standard Assessment Tasks and the Reliability of National Curriculum Assessment. *British Educational Research Journal*. 21, 2, 237-241.
- SALVIA, J. y YSELDYKE, J.E. (1995). *Assessment*. Boston: Houghton Mifflin Company.
- SCRIVEN, M. (1973). Goal-free evaluation. En House, E. (comp.): *School Evaluation: the politics and process*. Berkeley: Mc.
- SCRIVEN, M. (1995). A unified theory approach to teacher evaluation. *Studies in educational evaluation*. 21, 2, 111.
- SCRIVEN, M. (1994). Evaluation as a discipline. *Studies in Educational Evaluation*. Vol. 20, 147-166.
- SCHAFER, W.D. y LISSITZ, R.W. (1987). Measurement training for school personnel: Recommendations and reality. *Journal of Teacher Education*, 38 (3), 57-63.
- SCHÖN, D.A. (1992). *La Formación de Profesionales Reflexivos*. Madrid/Barcelona, Paidós/Mec.
- SECHREST, L. and FIGUEREDO, A.J. (1993). Program Evaluation. *Annual Review Psychology*, 4, 645-674.
- SECHREST, L. (1992). Roots: back to our first generations. *Evaluation Pract.* 13, 1-7.
- SERRANO, J.; TOJAR, J.C.; GALANTE, R. y ELÓSEGUI, E. (1995). Aportaciones recientes en medición y evaluación educativa. *Revista de Investigación Educativa*. n. 26, 129-158.
- SHAW, K.E. (1966). Cultural Issues in Evaluation Studies of Middle Eastern Higher Education. *Assessment & Evaluation in Higher Education*. 21, 4, 313-324.
- STAKE, R.E. (1983): La evaluación de programas; en especial la evaluación de réplica. En Dockrell, W. y Hamilton, D. (Eds.): *Nuevas reflexiones sobre la investigación educativa*. Madrid: Narcea.
- STAKE, R.E. (1967). The countenance of educational evaluation. *Teachers college record*. 68, 523-40.
- STENHOUSE, L. (1975). *An introduction to curriculum research and development*. Heinemann. London.
- STIGGINS, R.J. y FAIRES-CONKLIN, N. (1988). *Teacher training in assessment*. Portland, OR: Northwest Regional Educational Laboratory.
- STUART, M. (1995). Prediction and qualitative assessment of five-and six-year-old children's reading: a longitudinal study. *British journal of educational psychology*. 65, 3, 287.
- STUFFLEBEAM, D.L. (1995). Evaluation of superintendent performance: toward a general model. *Studies in educational evaluation*. 21, 2, 153.

- STUFFLEBEAM, D.L. (1971). *Educational evaluation and decision making*. Peacock. Itasca, III.
- SUMSION, J. and FLEET, A. (1966). Reflection: can we assess it? Should we assess it? *Assessment & Evaluation in Higher Education*. 21, 2, 121-130.
- TOOHEY, S.; RYAN, G. and HUGHES, C. (1996). Assessing the Practicum. *Assessment & Evaluation in Higher Education*, 21, 3, 215-229.
- TORRANCE, H. (1993). Formative assessment: some theoretical problems and empirical questions. *Cambridge Journal of Education*, 23 (3), 333-343.
- TUNSTALL, P. and GIPPS, C. (1995). *How does your teacher help you to make your work better? Children's Understanding of Formative Assessment*. Paper presented to the British Educational Research Association Conference, Bath UK.
- TYLER, L.E. (1950): *Test and measurements*. Prentice-hall. New York. (Trad. En Ediciones el Castillo, 1972).
- VIDICH, A.J. y LYMAN, S.M. (1994). Qualitative methods: Their history in sociology and anthropology. En *Handbook of Qualitative Research*, Dezin, N.K. y Lincoln, Y.S. (eds.).
- THOUSAND OAKS, CA. SAGE Publications, Inc., 23-59.
- WAKS, S. and BARAK, M. (1966). Role of evaluation in an interdisciplinary educational program. *Studies in Educational Evaluation*. 22, 2, 171-180.
- WEBSTER. W.J. (1995). The connection between personnel evaluation and school evaluation. *Studies in Educational Evaluation*. 21, 2, 227.
- WEINSTEIN, C.S. (1991). The classroom as a social context for learning. *Annu. Rev. Psychol.*, 42, 493-525.
- WESTON, C.; MCALPINE, L. y BORDONARO, T. (1995). A model for understanding formative evaluation in instructional design. *Educational technology research and development*. 43, 3, 29.
- WESTON, C. and McALPINE, L. (1994). *Improving Instruction Through Formative Evaluation*. A.E.R.A., New Orleans.
- WHYTE, W. F. (1943). *Street corner society: The social structure of an Italian slum*. Chicago: University of Chicago Press.
- WILIAN, D. and BLACK, P. (1966). Meaning and Consequences: a basis for distinguishing formative and sumative functions of assessment? *British Educational Research Journal*. 22, 5, 537-548.
- WILIAN, D. (1966). *National Curriculum Assessment and Programmes of Study: validity and impact*. *British Educational Research Journal*. 22, 1, 129-142.
- WORTHEN, B.R. & SANDERS, J.R. (1987): *Educational Evaluation. Alternative approaches and practical guidelines*. Longman Inc. London.
- WORTHEN, B.R.; BORG, W.R. y WHITE, K.R. (1993). *Measurement and Evaluation in Schools*. New York: Longman.
- ZABALA, A. (1966). La información de los procesos y los resultados del aprendizaje. *Aula de Innovación Educativa*, 55, 41-48.

LA EVALUACIÓN DE LOS ALUMNOS EN EL CONTEXTO DE LA EVALUACIÓN DE LA CALIDAD DE LAS UNIVERSIDADES

Fuensanta Hernández Pina
Universidad de Murcia

RESUMEN

La evaluación de los alumnos como un indicador de la calidad de las universidades se ha convertido en un tema de especial atención por parte de numerosos investigadores. En los modelos que se han elaborado desde planteamientos cualitativos-fenomenológicos se ha comprobado que la forma en que los profesores plantean la evaluación de los alumnos afecta a los enfoques de aprendizaje y a la calidad de dicho aprendizaje. Las conclusiones parecen apuntar a que si queremos mejorar el aprendizaje de los estudiantes hay que empezar cambiando las estrategias evaluativas dado el efecto que éstas pueden tener en la calidad total de dicho aprendizaje. Unas estrategias evaluativas cuantitativas llevan a enfoques superficiales de aprendizaje, mientras que la evaluación formativa y cualitativa puede producir enfoques de aprendizaje profundo y de alto rendimiento.

ABSTRACT

Students assessment as an indicator of the university quality is a topic of growing interest among many researchers. Students learning models from a qualitative-phenomenological point of view have shown that the way teachers assess students performance affects approaches to learning and the quality of that learning. It is a well established fact that in order to improve

student learning we have to change assessment strategies. Quantitative strategies of assessment will encourage surface approaches to learning whereas formative and qualitative strategies assessment encourage deep and achieving approaches to learning.

I. EL PLAN NACIONAL DE EVALUACIÓN DE LA CALIDAD DE LAS UNIVERSIDADES

La evaluación de la calidad de las instituciones universitarias es tema prioritario en muchos países desarrollados. Conocer las metas y objetivos que pretende una institución y evaluar en qué grado se han alcanzado, se ha convertido en el reto de aquellas instituciones que desean lograr prestigio y competitividad.

En España, al igual que en numerosos países de nuestro entorno, en el período de 1993-94, el Consejo de Universidades desarrolló un Programa para la Evaluación de la Calidad del Sistema Educativo en el que participaron varias universidades españolas. El Programa, al estar en fase experimental, no evaluó todas las titulaciones impartidas en cada universidad, sino solamente algunos títulos o aspectos determinados de los mismos.

El programa en esencia distinguía tres dimensiones en la evaluación: la enseñanza, la investigación y la administración de las universidades. En el nivel I se evaluó la totalidad de la Universidad y se tomaron en cuenta los indicadores cuantitativos, los análisis bibliométricos y los informes recogidos mediante cuestionarios. En el nivel II, se analizó la enseñanza y la investigación referidas a determinadas áreas de conocimientos. Los instrumentos utilizados para la recogida de los datos fueron muy similares a los utilizados en el nivel I, pero con un mayor grado de extensión. Al nivel III se incorporaron metodologías cualitativas para integrar en un único análisis los tres ámbitos en actuación de la Universidad mencionados más arriba.

Posteriormente, basándose en el Programa Experimental, el Consejo de Universidades, en su reunión de 25 de septiembre de 1995, acordó proponer al Gobierno un plan sistemático de evaluación de la calidad de las universidades españolas. Este plan que regula el modelo a seguir en dicha evaluación quedó aprobado en el Real Decreto 1947/1995, de 1 de diciembre, en el que se establece el Plan Nacional de Evaluación de la Calidad de las Universidades, con una duración de cinco años.

En el Plan se recogen todas aquellas dimensiones que los expertos han considerado básicas para la consecución de las metas y objetivos en una institución universitaria. Para Miller (1979) estos aspectos pueden sintetizarse en: metas y objetivos de la institución, el aprendizaje de los estudiantes, el rendimiento del profesorado, los programas académicos, los servicios, la gerencia, los aspectos financieros, el consejo de gobierno, las relaciones externas, el grado de preocupación por su mejora, etc.

La evaluación de todos estos aspectos puede abordarse desde dos perspectivas: a) desde la económica en la que los elementos contables y de control se convierte en aspectos prioritarios; y b) desde la académica o de control de los aspectos organizativos de los procesos de enseñanza y aprendizaje. Vroeijensti y Acherman (1991) han analizado ampliamente ambas dimensiones en un artículo titulado «Evaluación de la calidad basada en el control frente a la evaluación basada en la mejora» al que remitimos al lector.

La perspectiva que se adopte va a depender del concepto que se tenga de calidad —se evalúa para controlar o para mejorar— y de las relaciones que se establezcan entre los gobiernos y las instituciones. Una evaluación para el control de la institución por parte de los gobiernos apuntaría hacia un modelo de evaluación externa donde personas especializadas o mediante una comisión exterior a la universidad llevan a cabo la evaluación. Una evaluación para la mejora de la calidad llevaría a una evaluación interna, en la que los propios miembros de la institución serían los agentes de la evaluación. Ambos tipos de evaluación lejos de contraponerse se complementan puesto que hay aspectos que necesitan de ese control externo y otros que precisan de una autorregulación.

Con el fin de dar apoyo financiero al Plan Nacional de Evaluación, el 21 de febrero de 1996 aparece una Orden (B.O.E. de 28 de febrero de 1996), donde se convocan ayudas a proyectos de evaluación con los siguientes objetivos:

- a. «Promover la evaluación de la calidad de las Universidades españolas, tanto en el ámbito de la enseñanza, como de la investigación y demás servicios que la Universidad presta a la sociedad».
- b. «Facilitar a las Universidades y a las administraciones Educativas una metodología homogénea y unos criterios básicos comunes para la evaluación de la calidad compatibles con la práctica vigente en el contexto europeo».
- c. «Proporcionar a la sociedad, especialmente a los estudiantes universitarios, información relevante y objetiva sobre la calidad de las Universidades españolas, sus diferentes programas de estudio, sus áreas de especialización científica y el nivel de prestaciones y servicios que ofrece».
- d. «Proporcionar a las Administraciones Educativas y al Consejo de Universidades una información objetiva sobre el nivel de calidad alcanzado por las Universidades que pueda servir de base para la adopción de decisiones en el ámbito de las respectivas competencias».

El paso siguiente en la implantación de la evaluación por parte del Ministerio ha sido la elaboración de un documento-guía, enviado a todas las universidades, en el que se establecen de forma detallada los ámbitos de aplicación de dicha evaluación. Esta Guía, en su apartado II, incluye toda una serie de consideraciones y los protocolos para la evaluación de las tres dimensiones ante dichas: la enseñanza, la investigación y la gestión.

El apartado II.A, que es el que nos interesa destacar en este trabajo, dedicado a la evaluación de la enseñanza, contempla varios subapartados: la evaluación del contexto institucional; la evaluación de metas, objetivos y la planificación de la titulación objeto de estudio; la evaluación del programa de formación, en el que se incluyen aspectos como la estructura del plan de estudios, los programas de las asignaturas y la organización de la enseñanza; la evaluación del desarrollo de la enseñanza en la que se incluyen aspectos como la atención tutorial, la metodología docentes, **el trabajo de los alumnos, la evaluación de los aprendizajes y los resultados de la enseñanza**; en la evaluación de los alumnos recomienda evaluar las políticas de acceso, los servicios

específicos, etc.; en lo referente al profesorado incluye la distribución departamental de la docencia y la tipología del profesorado implicado en la docencia; en esta dimensión de la enseñanza se incluye la evaluación de las instalaciones y las relaciones externas. Como podemos observar, entre los aspectos que se recomienda evaluar dentro de la enseñanza aparecen tres epígrafes a mi juicio muy importantes pero insuficientemente tratados. Nos referimos al trabajo de los alumnos, la evaluación de los aprendizajes y los resultados de la enseñanza.

En **el trabajo de los alumnos** se hace referencia explícita a tiempo real invertido por los alumnos al estudio, preparación de trabajos, etc. Es decir, a las exigencias impuestas por el profesor en cada materia.

En **la evaluación de los alumnos** se aborda de una forma, en mi opinión, superficial el tema de la evaluación del aprendizaje, sugiriendo única y exclusivamente el análisis de los sistemas de evaluación utilizados por los profesores a partir de los programas hechos públicos por los departamentos o facultades, centrándose en la tipología y peso en la calificación así como su pertinencia en relación con los objetivos de la materia y el plan de estudio. Termina el apartado llamando la atención para que se «tomen en consideración los aspectos formales de garantía de la objetividad en la calificación», sin entrar en quién y cómo se garantizará tal objetividad.

En **los resultados del aprendizaje** se habla de dos tipos de resultados: inmediatos y diferidos. Dentro de los resultados inmediatos se ha incluido la tasa de abandonos, tasa de presentados a los exámenes, tasa de éxito, tasa de rendimiento, primera tasa de ingreso-egreso, segunda tasa de ingreso-egreso. En los resultados diferidos se incluye indicadores como los relacionados con la tasa de empleo a corto y medio plazo, niveles de sub-empleo, grado de satisfacción de graduados y empleadores, etc. Todos estos resultados, los inmediatos y los diferidos, serán recogidos en la evaluación rebasando la opinión del momento, tal y como se señala en la propia Guía.

Todas estas recomendaciones, en cada uno de los apartados, vienen acompañadas por unos protocolos en los que se incluyen preguntas que deseamos destacar:

«4.3 El trabajo de los alumnos

¿De qué manera se fomenta el aprendizaje independiente del alumno y su propia responsabilidad?

¿Se conoce la demanda del trabajo personal que ha de realizar el alumno individualmente?

¿Cuál es la valoración del Comité sobre el nivel real de exigencia en la titulación?

¿Cuál es la valoración sobre la factibilidad del Plan de Estudios desde la perspectiva del alumno (demanda de horas de clase, trabajo personal, número de años para finalizar la carrera, etc.)?

En relación a las cifras de matrícula, ¿Cuál es el porcentaje medio de asistencia a clase? Estímese en el período control de desarrollo de la asignatura, y analícese la información desagregada por ciclos, asignaturas obligatorias y optativas, primer año, distintos turnos si los hubiese, etc.

4.4 La evaluación de los aprendizajes

¿En qué momento y lugar se hace pública la convocatoria de los exámenes finales?

Analícese y valórese la práctica mayoritaria en relación con los siguientes aspectos (utilícese, si se cree oportuno, la tabla —cuadro de apoyo— para la recogida de la información correspondiente):

El tipo y el contenido de los exámenes

Los mecanismos de corrección

Los criterios que se utilizan para la evaluación

La periodicidad de la evaluación (finales, parciales, continuos, etc.)

¿Existe algún criterio para compensar o promediar notas en las diferentes asignaturas de un semestre o curso?

¿Cómo y cuándo se informa a los alumnos de los resultados de los exámenes parciales y finales?

¿Existe normativa específica a la revisión de exámenes?

4.5 Los resultados de la enseñanza

Analícese, por asignatura, los resultados académicos en primer y segunda convocatoria considerando los presentados y los aptos en relación con los matriculados.

Respecto al rendimiento académico global de la titulación acumúlase y valórese los siguientes indicadores para el último curso académico:

— Tasa de éxito

— Tasa de retraso

— Tasa de abandono

— Duración media de los estudios

Respecto al rendimiento en el transcurso de los estudios calcúlese y valórense los siguientes indicadores de rendimiento par el último curso académico:

— Tasa de abandono en el primer año

— Tasa de abandono en los dos primeros años

— Tasa de presentados

— Tasa de éxito

— Tasa de rendimiento

Respecto al rendimiento diferido analícese, si se dispone de información pertinente, los siguientes aspectos:

— El empleo y la demanda de los graduados de esa titulación

— La proporción de titulados que tienen su primer trabajo directamente relacionado con los estudios realizados

— Las opiniones de los titulados respecto a su formación, a la hora de buscar y/o encontrar empleo

— Las opiniones de los empresario respecto a la formación de los titulados que acceden al mundo laboral».

De nuevo observamos que la evaluación de los aprendizajes queda reducida a unos cuantos datos cuantitativos que, sin duda, son importantes pero que poco dicen acerca de la relación enseñanza-aprendizaje y la calidad del aprendizaje de los estudiantes.

2. MODELO DE EVALUACIÓN DE LA ENSEÑANZA

La enseñanza junto con la investigación constituyen las primeras funciones de la universidad. En un informe elaborado por Atkins, Beattie y Dockrell (1993) sobre evaluación de los aprendizajes en educación superior abordan el tema de cuales deberían ser los propósitos más importante de las instituciones universitarias. Estos autores sugieren que tales objetivos pueden concretarse en cuatro apartados:

- a. Facilitar una experiencia educativa general
- b. Preparar a los alumnos para el conocimiento creativo, su aplicación y difusión
- c. Preparar a los alumnos para una profesión específica u ocupación profesional
- d. Preparar para el empleo en general.

Estos objetivos lejos de contraponerse se complementan y pueden convertirse en los puntos fuertes en la formación de un alumno, el contenido de un curso o de una institución. De ahí que hayan sido desarrollados en sub-componentes o competencias más específicas, como son:

- a. Para el objetivo primero: el desarrollo de una mente entrenada, la adquisición de conocimientos para ser una persona educada, el desarrollo personal para la vida adulta y el establecimiento de las bases para un aprendizaje continuo.
- b. Para el objetivo segundo las competencias podrían ser: adquirir el marco conceptual de las materias de la titulación; conocer con profundidad algunos aspectos de las materias con vista a una especialización; comprender la metodología y procedimientos de investigación específicos de las disciplinas; desarrollar proyectos e investigaciones a pequeña escala; desarrollar la solución de problemas; desarrollo de técnicas de comunicación; desarrollar técnicas de grupo, etc. Todo esto dentro de un marco de aprendizaje experiencial.
- c. Para el tercer objetivo proponen las siguientes subcategorías o competencias: integración de la teoría y la práctica; adquisición de experiencias derivadas del conocimiento y de su aplicación al contexto; desarrollo de habilidades y competencias profesionales derivadas de experiencias profesionales de primera mano; adquisición de normas, actitudes, cualidades personales específicas de la profesión; comprender y adquirir las competencias éticas de la profesión; compren-

- der el contexto organizacional de la profesión; desarrollar la habilidad de reflejar la propia práctica, autoevaluarse, desarrollarse profesionalmente, etc.
- d. Para el cuarto objetivo, la preparación para el empleo, las competencias podrían sintetizarse en: trabajar en contexto diversos; desarrollar la habilidad de reflexionar y aprender de la experiencia; desarrollar habilidades diversas como, por ejemplo, las comunicativas, informáticas, lenguas extranjeras, automotivación, liderazgo, trabajo en grupo, solución de problemas, evaluación de riesgos y sus consecuencias, etc.

Estos son a grandes rasgos los conocimientos, comprensión y habilidades que los estudiantes deben adquirir a su paso por las instituciones universitarias. Lo que los profesores evalúen y cómo lo evalúen va a afectar a la calidad de tales aprendizajes. Si sólo evalúan niveles inferiores del aprendizaje, esto indudablemente afectará a los procesos de aprendizaje desarrollados por los propios estudiantes.

Todos los componentes examinados más arriba aparecen relacionados con la enseñanza que los profesores deben ejercer en la universidad. Efectivamente la relación entre lo que el profesor enseña y evalúa, y lo que el alumno aprende, debe abordarse de forma prioritaria, entendemos, en la evaluación institucional; concretamente en la evaluación de la enseñanza como elemento a destacar en la producción de una institución universitaria.

Tanto Rodríguez Espinar (1991) como De Miguel (1991) señalan que conviene evitar el papel reduccionista que suele asignarse a la evaluación de la enseñanza; ambos autores destacan que la evaluación de la enseñanza universitaria no debe en ningún momento centrarse exclusivamente en el profesor, sino más bien en todos aquellos elementos que rodean la actividad que el profesor lleva a cabo en el aula.

Rodríguez Espinar (1991) considera que la enseñanza universitaria ha de entenderse como «el conjunto de actividades y servicios desarrollados por una institución universitaria y orientados a la formación de graduados» señalando la docencia como un componente de tal conjunto. De Miguel (1991), por su parte, distingue entre lo que debe entenderse por una unidad de docencia y la actividad docente que realiza un profesor. De ahí que abogue por evitar el reduccionismo de evaluar sólo al profesor como único elemento de la evaluación de la unidad docente. De Miguel, además, es partidario de plantear la evaluación de la enseñanza con un enfoque de innovación y cambio de los procesos de enseñanza destacando el carácter positivo y de mejora que ha de tener la evaluación.

Este mismo autor destaca cuales son los componentes principales de la enseñanza o docencia que ayudan a evitar dicho reduccionismo. Para ello aparte del principio de que «la dimensión docente de la universidad no se limita a la enseñanza que realiza un profesor concreto en su aula, sino que implica la actuación de una serie de órganos e individuos que toman decisiones sobre qué cosas se van a enseñar, cómo, cuándo y por quiénes van a ser impartidas las enseñanzas y con qué objetivos o finalidad».

De Miguel (1991) es partidario de un modelo de evaluación basado en la producción como es el modelo de Stufflebean. En la universidad se inicia la tarea cuando se establecen las ofertas de títulos y cursos y los alumnos (input) con las características que demandan tales titulaciones. A continuación la institución organiza las enseñan-

zas y las actividades propias con el fin de que aquellos adquieran los conocimientos y habilidades (procesos) propios de la titulación. Por último, como resultado de tales procesos se obtienen unos productos reflejados en el alumno y en la institución. Estas tres fases se efectúan dentro de un contexto que va a ser específico a cada institución y que será la que mediatice y envuelva no solo a las concepciones que la institución tenga sobre la enseñanza sino al propio proceso de producción (contexto).

En cada una de estas fases (contexto, input, proceso y producto) se dan niveles de administración diversos, con unidades de gestión y niveles de responsabilidad propios que se pueden estar anidados unos en otros. Dado que cada uno de ellos tiene responsabilidades propias, cualquier modelo de evaluación debe evaluarlas desde las funciones que se le hayan asignado y desde su propia responsabilidad. De Miguel diferencia tres niveles de competencias: unas que emanan de los órganos de gobierno de la institución, otras que emanan de unidades más pequeñas como pueden ser los centros, departamentos, etc. y, por último, las que emanan del profesor individualmente. En un trabajo posterior, De Miguel (1995) llama la atención sobre el modelo de producción contexto/recursos/proceso/producto señalando que este modelo puede aplicarse a los distintos niveles del sistema educativo. El modelo, siendo el mismo, cambia en los factores o variables que incluye en cada caso. Sin embargo, la calidad obtenida en cada nivel va a depender del nivel previo, a la vez que influye en el nivel siguiente. Ello significa que «la política educativa a nivel estatal constituye el referente que enmarca las actividades que desarrollan las instituciones, como a su vez éstas determinan la calidad de la enseñanza en las aulas» (p. 37).

Para Biggs (1996) la evaluación del aprendizaje ocupa un lugar clave en la determinación de los resultados del aprendizaje. Esto significa que la evaluación no es un hecho aislado en una institución, ya que ésta es un sistema holístico que por su estructura organizativa incluye niveles y funciones que determinan los procedimientos de enseñanza y evaluación y que, a su vez, incide en la percepción que los alumnos tienen sobre qué y cómo aprender. Esta es, a juicio de Biggs, la razón más importante de por qué cada institución superior debería tener su propia política y guía sobre la evaluación, suministrando una serie de principios coherentes y conocimientos procedimentales acerca de la evaluación. Dicha política permitiría asegurar la coherencia entre los propósitos y objetivos de la institución, y por extensión de las titulaciones que imparte, y la evaluación de los logros alcanzados tras los procesos de enseñanza-aprendizaje. Biggs ilustra esta idea conforme vemos en la figura 1 en la que observamos que el sistema de evaluación (qué es probable) implica un compromiso entre los objetivos (qué queremos), cómo lograr tales objetivos (qué es posible) y lo que la administración y nuestros colegas nos permiten (qué es permitido). Los objetivos van encaminados a la toma de decisiones en primer lugar. La tecnología nos permitirá juzgar si tales objetivos se han alcanzado. A continuación nos encontramos con los requisitos formales e informales del sistema social. Todos estos componentes pueden afectar a los anteriores, es decir, podríamos cambiar nuestros objetivos a la vista de lo posible o lo permisible. El sistema social podría impedir algunas cosas que son técnicamente posibles; podría incluso ser la medida de lo deseable y de los cambios posibles que son permitidos.

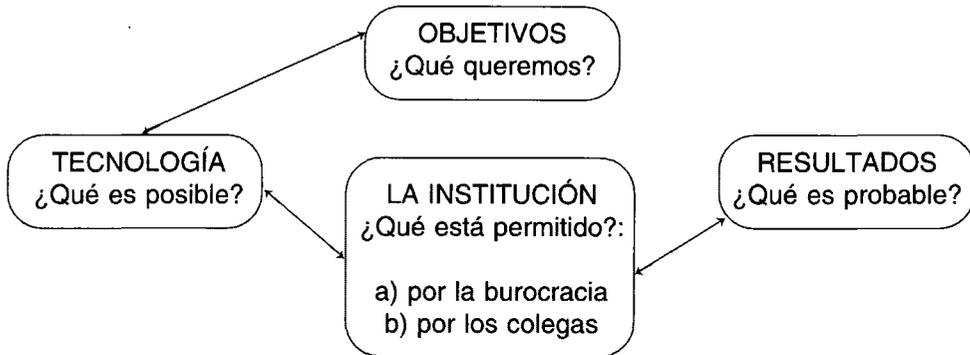


Figura 1

3. UN MODELO DE APRENDIZAJE DE CALIDAD

Desde la IV Conferencia Internacional sobre Educación Superior celebrada en Lancaster en 1978 el estudio del aprendizaje de los alumnos universitarios ha experimentado un cambio de estrategia pasando de un análisis exclusivamente cuantitativo a otro más cualitativo donde los aspectos contextuales han adquirido un gran protagonismo para comprender cómo los alumnos abordan su aprendizaje y estudio. Housell, Marton Säljö, Entwistle, Biggs, Meyer, Muller, etc. han sido algunos de los autores que más han contribuido a enriquecer esta nueva perspectiva en el estudio del aprendizaje universitario y a establecer qué variables son las más relevantes para la calidad de dicho aprendizaje. Entre ellas cabe destacar, sin lugar a dudas, los sistemas de evaluación adoptados por los profesores. El efecto retroactivo de la evaluación determina lo que el alumno aprende y cómo lo aprende más que el propio currículum y los métodos de enseñanza utilizados por los profesores. Si la evaluación tiene unos efectos tan importantes sobre el aprendizaje de los alumnos estamos de acuerdo con autores como Elton y Laurillard (1979), Biggs (1996) en preguntarnos por qué no mejorar el aprendizaje de los alumnos cambiando simplemente los sistemas de evaluación. La respuesta a esta pregunta es un tanto compleja y no tiene solución única y aislada. Crooks (1988) en la revisión que hace del impacto de la evaluación en los estudiantes apunta algunas respuestas que comentaremos más abajo.

Perspectivas de la investigación sobre el aprendizaje de los estudiantes universitarios

Antes de entrar a describir brevemente las dos concepciones que orientan la forma de abordar el aprendizaje de los estudiantes vamos a hacer un breve comentario sobre la terminología al uso. Esta nomenclatura ha ido cambiando tanto en una como en otra perspectiva conforme las investigaciones se han ido sucediendo.

Marton y Svensson fueron los que inician en 1979 una actividad investigadora que les llevaría a establecer una distinción entre un enfoque **cuantitativo** o experimentalis-

ta (el único hasta entonces en vigor) y un enfoque **cuantitativo** u observacional que se situaba más en la línea interpretativa. Basándose en la técnica de la entrevista hecha a los estudiantes, realizaron por primera vez una descripción de los diferentes enfoques usados por los estudiantes para abordar su aprendizaje y estudio desde supuestos netamente cualitativos.

Van Rossun y Schenk (1984) han hablado de dos perspectivas: una de **primer orden** y otra de **segundo orden** (esta segunda etiqueta tomada de Marton, 1981). La perspectiva de primer orden pondría el énfasis en la *descripción* de aspectos de la realidad que influye en el aprendizaje, tales como el comportamiento de los estudiantes, el ambiente de aprendizaje, la evaluación etc. La perspectiva de segundo orden incidiría más en tratar de comprender el modo en que el alumno *percibe* dicha realidad. Realidad que no es descrita por el investigador tal y como él la percibe, sino desde la óptica del propio estudiante. Marton (1981) ha denominado a este tipo de método 'fenomenografía'.

Ramsden (1985), por su lado, distingue entre aquellas investigaciones que tienen un carácter **nomotético** y las que tienen un carácter **idiográfico**. La perspectiva nomotética trataría de derivar leyes sobre el aprendizaje para descubrir las características que determinan el progreso de los alumnos. La investigación idiográfica, por el contrario, se centraría en la experiencia única de los estudiantes y en la variedad de cualidades que aquellos despliegan.

El rasgo que parece caracterizar a muchas de las investigaciones dentro de esta perspectiva es que la calidad del aprendizaje de los estudiantes depende del **enfoque** adoptado para aprender. Este enfoque a su vez va a depender de las experiencias previas y del modo como interprete el alumno los requisitos del **contexto del aprendizaje**, un contexto en el que ciertos aspectos del mismo pueden ser seleccionados por el estudiante (materias optativas), pero no así otros (tipo de enseñanza, formas de evaluación etc.). La percepción o idea que los estudiantes se formen de lo que el contexto les exige y les ofrece es lo que en última instancia determinará su *modo* de abordar las tareas de aprendizaje en un sentido determinado y su *enfoque* de las mismas.

Esta secuencia la esquematiza Ramsden del modo siguiente (figura 2).

Entwistle (1992) ha hecho también una doble distinción. La primera perspectiva **cuantitativa**, deriva de la psicología cognitiva, y estaría en línea con los planteamientos cuantitativos y experimentalistas. La segunda más **cuantitativa**, está basada en la entrevista y la observación de los contextos naturales. El objetivo estriba en identificar conceptos y categorías que derivan de las actividades del estudio diario que los estudiantes dicen utilizar. Cole (1990) hace también una diferenciación en las concepciones en el rendimiento o logros en el aprendizaje: la caracterizada por el desarrollo de habilidades y el aprendizaje de hechos y la caracterizada por el desarrollo de habilidades superiores y el aprendizaje de conocimiento avanzados. Esta dos concepciones tienen su impacto en las investigaciones que se realizan, en el modo en que los profesores llevan a cabo su enseñanza y en el diseño de los metas y objetivos institucionales. La autora propone que ninguna de esta dos concepciones es suficiente para orientar el pensamiento en el ámbito educativo. A cambio sugiere que habría que buscar un marco más amplio que permitiese comprender las dos concepciones por

separado y en relación la una con la otra; atender la implicaciones que para la instrucción y el currículum tiene la concepción que adoptemos sobre el aprendizaje o los logros; y, por último, establecer una relación más clara de nuestra concepción sobre el aprendizaje o los logros con las metas y propósitos a corto, medio y largo plazo. Este último requisito se relaciona con lo que queremos que nuestros alumnos aprendan hoy, mañana y lo que queremos que recuerden y sean capaces de hacer de aquí en adelante.

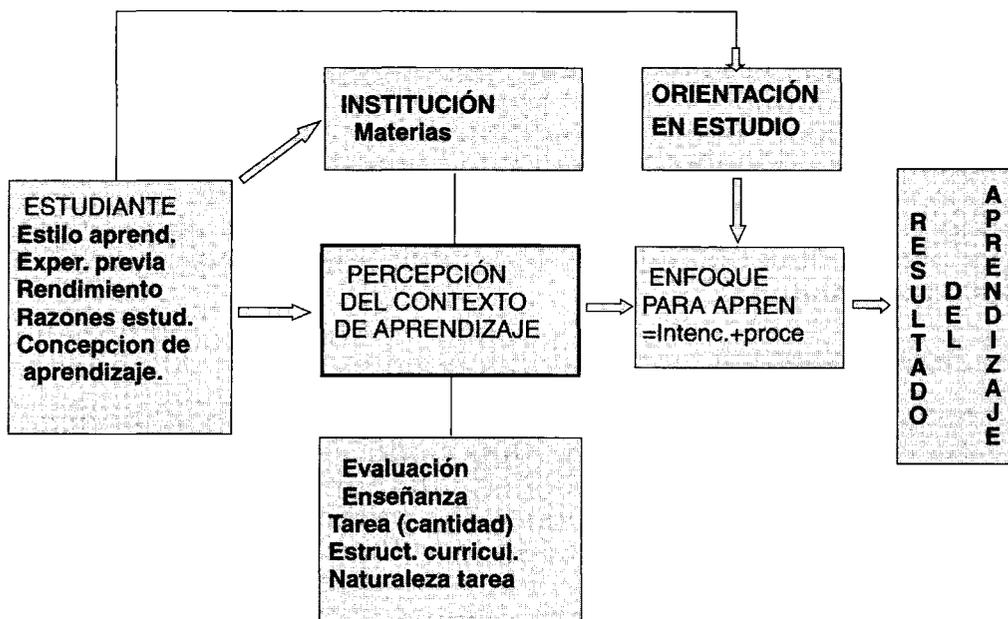


Figura 2.

Comprendiendo el aprendizaje del estudiante (Ramsden, 1985).

De la revisión hecha, observamos que todos los autores vienen a coincidir en dos paradigmas de investigación en el estudio del aprendizaje de los alumnos, coincidentes a su vez con las dos perspectivas de investigación existentes en el campo de la investigación educativa. Ambas perspectivas, lejos de convertirse en posturas enfrentadas, han propiciado hallazgos complementarios, pese a que sigan existiendo discrepancias respecto a los tipos de intervención que se debería seguir para mejorar los enfoques de aprendizaje (Entwistle, 1992).

a. La orientación cuantitativa

Veamos a continuación de un modo más detenido las aportaciones que las dos perspectivas han realizado, empezando por la orientación cuantitativa en la que cabe distinguir dos líneas fundamentales de acción: la **conductista** y la **cognitiva**.

a.1. *Perspectiva conductista*

Esta perspectiva concibe el aprendizaje de los alumnos como la acumulación de contenidos aprendidos; un buen estudiante es aquel que más sabe. Desde sus orígenes, la enseñanza y la investigación sobre el aprendizaje de los alumnos estuvieron influidas por el neo-conductismo siendo tema casi exclusivo de la psicología experimental. Esta perspectiva del aprendizaje ha dado lugar a un corpus teórico de innegable valor al abordar temas tales como la generalización, la discriminación, la retención, la transferencia, los conceptos de aprendizaje, etc., siendo el hilo conductor de análisis las relaciones entre estímulo y respuesta.

La tradición conductista incorporó términos tales como '**hábitos de estudio**' prestando más interés a la adquisición y entrenamiento de habilidades fácilmente identificables así como a los medios y las técnicas para su desarrollo y entrenamiento. Esta perspectiva, conocida dentro de la temática que nos concierne también como '**entrenamiento en habilidades para el estudio**' o '**aprendizaje de técnicas de estudio**' persigue, pues, el desarrollo de habilidades en relativo aislamiento del contexto y de los objetivos del aprendizaje.

a.2. *Perspectiva cognitivista*

En las últimas décadas se ha desarrollado una teoría del aprendizaje basada en el análisis detallado de la adquisición de estructuras del conocimiento a través del uso de lenguajes de programación, lo que ha permitido plantear nuevas hipótesis acerca de los procesos cognitivos y estructuras intervinientes tanto en la solución de problemas como en la comprensión del lenguaje. La llegada de la psicología cognitiva ha supuesto un cambio paradigmático en la investigación del aprendizaje desbancando de este modo la hegemonía que tuviera el neo-conductismo.

El cognitivismo asigna mayor protagonismo al sujeto en el aprendizaje. Considera al sujeto un participante activo en el proceso enseñanza-aprendizaje, de tal forma que el efecto de la enseñanza va a depender —al menos parcialmente— de lo que el estudiante conoce (sus conocimientos previos) y lo que hace durante dicho aprendizaje (Meyer, 1983, 1984; Dansereau, 1985; Winstein, 1978; Wittrock, 1978), etc. El objetivo va ser identificar todas aquellas estrategias que se pueden enseñar a un alumno para que las utilice durante su aprendizaje. Estas estrategias de aprendizaje se han definido como comportamientos y pensamientos del alumno durante el aprendizaje, y tienen por finalidad influir en el proceso de codificación del alumno.

Esta nueva visión del aprendizaje, ha llevado a un cambio de visión en las formas de enseñar. El profesor no sólo ha de preocuparse por el producto del aprendizaje, (es decir, **qué** debería saber el alumno como resultado del aprendizaje), sino que ha de tener en cuenta los procesos que intervienen en el aprendizaje, debiendo enseñar a los alumnos técnicas y estrategias que les faciliten dicho aprendizaje. Es decir, ha de enseñar **cómo aprender**.

Las *estrategias de aprendizaje* es lo que distancia a la teoría cognitiva de la teoría conductista, pues la primera busca comprender cómo la información que entra en el

sujeto es procesada y estructurada en la memoria, convirtiendo de este modo el aprendizaje en un proceso activo que tiene lugar dentro del estudiante y que éste puede alterar. De este modo, el aprendizaje ya no es el resultado directo de la instrucción, sino que va a depender tanto de la presentación que haga el profesor de la materia como del modo de procesar el alumno dicha información (Hernández Pina, 1993).

La evaluación del aprendizaje dentro de esta orientación supone la aplicación de pruebas a los alumnos que expresen la cantidad y corrección de sus respuestas de acuerdo con el instrumento aplicado. La nota obtenida es la suma de todas las competencias mostradas en el examen. Para Biggs (1996) esta tecnología puede ser apropiada en aquellos casos en los que el aprendizaje implica una agregación de cuestiones aprendidas con cierta exactitud (como por ejemplo, fórmulas, vocabulario, etc.), pero puede ser un error hacer una generalización implícita de que todo el aprendizaje se adquiere en porciones unitarias y pueden ser evaluadas del mismo modo.

b. La orientación cualitativa

La orientación cualitativa ve el aprendizaje no sólo como la acumulación de conocimiento, sino como la interpretación de tales conocimientos y la incorporación del nuevo material a los conocimientos previos. La perspectiva cualitativa intenta describir los procesos de estudio de forma holística, aunque también se han hecho esfuerzos por identificar conceptos y categorías con el fin de poder transmitir mejor los resultados de la investigación. El corpus conceptual que se ha desarrollado, se ha obtenido fundamentalmente a través de la experiencia directa con los propios estudiantes, vía entrevista. Dichos conceptos forman una jerarquía que va desde un nivel amplio e inclusivo a aquel que describe las reacciones a tareas específicas. Considerados globalmente, muestran una coherencia que es útil a la hora de diseñar programas de intervención para la mejora de la calidad del aprendizaje de los alumnos.

En esta orientación cualitativa podemos diferenciar dos líneas de investigación claramente definidas: la que se interesa por el estudio de los **estilos** de aprendizaje, y la que se centra en los **enfoques** de aprendizaje. El análisis de estos términos ha sido abordada en otro lugar al que remitimos al lector (Hernández Pina, 1993).

La evaluación dentro de esta orientación cualitativa tiene un carácter longitudinal. Los resultados del aprendizaje se convierten en construcciones que el estudiante ha ido realizando en cada etapa de su aprendizaje. La evaluación puede tener un doble carácter, como un proceso de desarrollo, es decir, el objetivo del aprendizaje es descubrir en qué fase se encuentra el sujeto y de esta manera orientarlo; o en un proceso ecológico, es decir, en qué medida el alumno aplica sus conocimientos a la solución de los problemas reales (Biggs, 1996). La evaluación centrada en el desarrollo es más apropiada para evaluar el aprendizaje de los conocimientos declarativos y la evaluación ecológica para evaluar los conocimientos procedimentales. La primera se centra en la comprensión de mundo y la segunda en cómo operar sobre él. El valor de la evaluación ecológica es que obliga a evaluar la preparación profesional y no sólo los conocimientos teóricos alcanzados en la institución.

b.1. Los enfoques de aprendizaje

Los enfoques de aprendizaje constituyen la perspectiva más reciente de la orientación cualitativa. En líneas generales se sitúa dentro del paradigma del procesamiento de la información, aunque con un planteamiento fenomenológico distinto. Dentro de esta perspectiva encontramos los trabajos de Marton, Säljö, Svensson, Entwistle, Ramsden, Watkins, Biggs, etc.

Para Marton y Svensson (1979) el aprendizaje depende de tres dimensiones: **a)** la conciencia que el estudiante tiene del aprendizaje, **b)** el contenido de las materias, y **c)** las demandas del contexto de aprendizaje. La **conciencia del acto del aprendizaje** es un aspecto básico en el estudio de los enfoques de aprendizaje. El **contenido** obliga al estudiante a adoptar estrategias distintas. A nivel universitario el contexto engloba variables tales como el sistema de evaluación, el estilo de enseñanza, las expectativas, las experiencias de aprendizaje previas, los tipos de materiales de aprendizaje, etc. Marton y Saljo (1976b) fueron los iniciadores del estudio de la influencia que estos factores contextuales tienen en el aprendizaje. Con sus investigaciones demostraron que las preguntas que se introducían en distintas tareas de aprendizaje ejercían un influjo considerable en cómo el alumno priorizaba la información.

Entwistle (1992) ha observado igualmente que en la universidad existe la tendencia a esperar que los estudiantes tengan metas académicas. Sin embargo, cuando los alumnos ha sido entrevistados, se ha observado que hay una gran diversidad en cuanto a los objetivos a alcanzar. El conocimiento de tal diversidad sirve para entender la forma en que los estudiantes se enfrentan al estudio. En un estudio, Taylor (citado por Gibbs y col., 1984), utilizando la técnica de la entrevista encontró cuatro orientaciones de carácter educativo, entendiendo por 'orientación' los motivos y propósitos que describen las actitudes de los estudiantes hacia los estudios que están realizando: *académica, vocacional, personal y social*.

b.2. Los enfoques 'profundo' y 'superficial' de aprendizaje. Un primer análisis

El estudio del aprendizaje de los estudiantes desde una perspectiva de segundo orden o fenomenológica ha recibido especial atención por parte del Grupo de Gotenburgo —Marton, Säljö, Svensson, etc. Estos autores han realizado un análisis cualitativo— naturalista en el que se pedía a los estudiantes que tras leer un texto, respondieran a una serie de preguntas. En las entrevistas Marton y Säljö encontraron que los estudiantes diferían notablemente en el nivel de comprensión de lo leído como consecuencia de lo que ellos empezaron a denominar 'enfoques de aprendizaje'. Observaron que los alumnos adoptaban un enfoque **profundo** o **superficial** en la lectura que realizaban de tales artículos. Estos términos que usaron estos autores para describir ambas formas de abordar las lecturas se generalizaron, sirviendo para describir el modo de abordar los alumnos no sólo la lectura, sino otras muchas tareas académicas: las clases, los trabajos que escriben, la solución de problemas, etc.

Las características que definen a los enfoques tienen su base en las intenciones mostradas por los estudiantes. De este modo, el **enfoque profundo** depende de la

intención que el sujeto tiene de alcanzar una comprensión personal de los materiales que estudia. Este enfoque parece tener sus raíces en una orientación de carácter educativo intrínseco y una concepción sofisticada del aprendizaje. Los que lo adoptan parecen interaccionar críticamente con el contenido, relacionándolo con sus conocimientos previos y evaluando los pasos lógicos que llevan a una determinada conclusión.

El **enfoque superficial**, en cambio, parece derivar de una orientación extrínseca y una concepción simplista del aprendizaje como es la memorización. Implica una intención de satisfacer las tareas o los requisitos del curso, vistas como imposiciones externas distantes de sus intereses personales. No es que este enfoque no pueda ser activo, como el anterior, pero busca más la identificación de aquellos elementos de la tarea más susceptibles de evaluación para memorizarlos.

Casi en paralelo a los trabajos de Marton y sus colegas, a finales de los años setenta y comienzos de los ochenta comienzan a aparecer otros trabajos sobre modelos de aprendizaje dirigidos por N. Entwistle y Ramsden, Biggs, etc. El objetivo principal de su línea de investigación era establecer cómo aprenden los estudiantes a partir de la información que ellos mismos suministran. Estos autores han llegado a conclusiones muy similares a las alcanzadas por Marton y su grupo y han desarrollado modelos de aprendizaje en los que incluyen las variables que según su teoría y los resultados de sus investigaciones determinan la calidad del aprendizaje. Por motivos de espacio solo incluimos el modelo de Biggs (1993) que pasamos a comentar brevemente.

En 1989 tras diversas investigaciones elabora Biggs un modelo de aprendizaje denominado 'Modelo 3P', por contemplar en él el contexto de enseñanza, los enfoques de aprendizaje y los resultados del aprendizaje formando un sistema en estado de equilibrio. En los trabajos de 1989 y 1990 ha dedicado Biggs una atención especial al papel que desempeñan los modelos de enseñanza en la configuración de los enfoques. Al igual que existe tres concepciones sobre el aprendizaje, al que tanto los alumnos como los profesores pueden sentirse atraídos, también existen básicamente tres concepciones de enseñanza: la *cuantitativa*, la *estratégica* y la *cuantitativa*.

Estas tres concepciones de la enseñanza presentan una estructura jerárquica, siendo las dos primeras las más socorridas a nivel universitario. Según un estudio llevado a cabo por Radloff y Sampson (1988) citado por Biggs (1989), la lección magistral sería la metodología de enseñanza más utilizada por los profesores universitarios.

A cada una de estas tres concepciones de la enseñanza subyace una teoría distinta sobre cómo se desarrolla la competencia en los alumnos. Así el modelo cuantitativo asume que el conocimiento existe en torno a uno a la espera de ser descubierto por alguien inteligente o que se encuentra en el momento oportuno en el sitio adecuado. Una vez descubierto, puede enseñarse en cantidades medibles por procesos de transmisión. Es así como la exposición se convierte en el modo más apropiado de enseñanza y de acumular conocimiento sobre los hechos. Dicho conocimiento, sin embargo, no altera en sí mismo la comprensión e interpretación que el sujeto tiene de los hechos.

En un trabajo publicado por Biggs titulado 'What do inventories of students' learning processes really measure? A theoretical review and clarification' (1993) incluye el modelo de aprendizaje '3P' (Figura 3) que se caracteriza por englobar tres

factores: presagio-proceso-producto, si bien con una relación recíproca entre todos sus componentes, convirtiendo de este modo el modelo en un sistema interactivo y complejo. Este modelo causal necesita de un análisis profundo desde los planteamientos de los modelos estructurales causales con el fin de poder determinar de una forma clara e integral cuál es realmente la estructura de la relaciones de los elementos que progresivamente se han ido integrando en él.

Los factores de *presagio* incluyen variables relacionadas con el alumno como son los conocimientos previos, las habilidades, el modo de aprendizaje preferido, los valores y las expectativas. La variables relacionadas con el contexto de la enseñanza incluyen toda una serie de superestructuras relacionadas con la institución y el profesor: la estructura del curso, el contenido curricular, los métodos de enseñanza, el clima de la clase y la evaluación. Estos dos bloques de variables se relacionan directamente con la naturaleza del resultado del aprendizaje y a través de las variables de proceso.

La novedad en esta relación es su simetría frente a la asimetría de modelos previos. De igual forma, los dos bloques de variables de presagio que en los modelos anteriores formaban dos grupos no relacionados entre sí, ahora aparecen simétricamente relacionados. La percepción que los profesores tienen sobre los motivos o habilidades de sus alumnos no figuraba en los modelos previos, cuando dicha percepción influye en las decisiones que tome el profesor sobre el tipo de enseñanza que imparte. Paralelamente, las percepciones que los estudiantes tienen acerca del contexto de la enseñanza, que tampoco recogían otros modelos, se incluyen ahora por entender que estas afectan directamente a sus motivos, predisposiciones y toma de decisión.

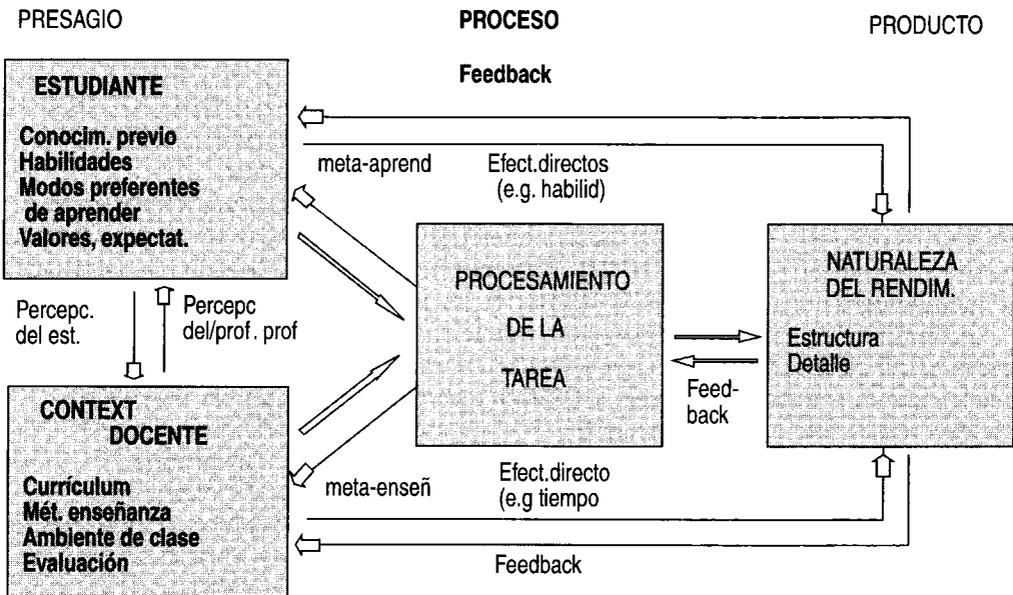


Figura 3.
Modelo '3P' (1993).

Esta percepción del sujeto que Biggs (1985, 1993) denomina meta-aprendizaje se centra en los procesos (en cómo realizar la tarea) en paralelo con el acto cognitivo de ocuparse o procesar el contenido del aprendizaje. De este modo el proceso tiene ahora dos sentidos totalmente distintos: el sentido meta-cognitivo (decidir cómo manejar la tarea en un determinado contexto) como una estrategia general, y el significado táctico que se refiere al tipo de procesos cognitivos (de nivel alto o bajo) que se usan. La idea de unir todos los componentes del modelo recíprocamente sirve para que todo el modelo se comporte como un sistema integral, de tal forma que el cambio en uno de sus elementos afectará a los otros, tal como sucede con los sistemas abiertos como por ejemplo la enseñanza universitaria. En caso contrario, el elemento revertirá a su status quo tal como sucede en los sistemas cerrados.

El modelo 3P (1993) se hace eco de distintos sistemas anidados que son relevantes en el aprendizaje de los estudiantes. Biggs menciona los cuatro siguientes:

- a) El *sistema del estudiante*, que comprende una relación equilibrada entre factores cognitivos y afectivos y un espacio fenomenológico percibido.
- b) El *sistema de la clase*, que abarca a los estudiantes, a los profesores y al marco docente.
- c) El *sistema institucional*, que contiene a su vez subsistemas como son los Departamentos y las Facultades, cada uno con su propio potencial para realizar o impedir una práctica informada.
- d) El *sistema comunitario*, que puede poner limitaciones en los distintos tramos del desarrollo educativo.

Cada uno de estos sistemas intenta un estado de equilibrio no sólo entre sus propios componentes, sino también con un modelo más general, lo cual es fundamental para saber si una determinada intervención es probable que produzca un cambio bien sea a nivel de estudiante, a nivel del docente o incluso a nivel de la propia institución.

La idea de una situación o estado estable es también de una gran ayuda para comprender cómo se han de usar los cuestionarios sobre los procesos de aprendizaje/ estudio para medir la calidad del aprendizaje que tiene lugar en el aula tanto a nivel individual como de grupo (Biggs, 1987a, 1987b). La predisposición que el sujeto tenga hacia un determinado enfoque de aprendizaje será el modo de que el estudiante logre el equilibrio en el sistema educativo tal como lo percibe. Es decir, dadas unas metas que el estudiante ha de alcanzar, la autopercepción de su habilidad, el modo de enseñanza y la evaluación, los resultados obtenidos, etc., le servirán para que, tras un período de exposición a un marco de enseñanza/aprendizaje concreto, desarrolle un determinado enfoque (superficial, profundo, etc.) que le permita llevar a cabo sus tareas académicas lo más cómodamente posible.

b.3. Aspectos que determinan los enfoques de aprendizaje

La comprensión de las razones de por qué un alumno estudia de un modo determi-

nado nos remite a considerar a los enfoques de aprendizaje en relación con el **contexto** académico o con el ambiente que rodea a las experiencias de aprendizaje. El modo de abordar el aprendizaje, como señalábamos, depende de la percepción que el sujeto tenga de lo que se le exige o de lo que considere decisivo en términos de la evaluación (Becker y col. 1968). Si muchos de los trabajos realizados durante la década de los ochenta se centraron en describir las diferencias cualitativas sobre la forma en que los alumnos abordan su trabajo académico, ahora el interés se ha desplazado hacia el estudio de los aspectos contextuales que influyen en dichos procesos de aprendizaje, tales como los métodos de enseñanza, el currículum, los procedimientos de evaluación, etc. Entwistle y Ramsden (1983), Marton y col. (1984), Thomas y Bain (1984), Crooks (1988), Entwistle y Tait (1990), Fransman (1991), Banta (1997), Biggs (1996), De Miguel (1996), Kniveton (1996), Hager y Butler (1996), Nuutinen (1997) etc.

Algunos investigadores han explorado la influencia que tiene la percepción que el alumno posee del contexto académico sobre los enfoques de aprendizaje y de determinadas materias (De la Orden, 1986). Una de las variables que más parece influir en la configuración de un enfoque dado son los procedimientos de evaluación: cómo los alumnos perciben dicha evaluación. De la Orden, en la revisión que hace de la «Investigación sobre la evaluación educativa» llama la atención sobre esta influencia al formular lo que en sus propias palabras podría ser la tesis general: «Sea cual fuere la modalidad, sistema o tipo de evaluación vigente en un contexto educacional, *su influencia sobre el proceso y el producto* (subrayado nuestro) de la educación es decisiva y afecta directamente o indirectamente a la calidad» (1983). Así, si por ejemplo las preguntas son tipo ensayo, la solución de problemas, los trabajos de investigación, etc., promoverán un aprendizaje profundo, mientras que el uso de pruebas objetivas inducirá a un enfoque superficial.

Entwistle y Ramsden (1983) diseñaron un cuestionario para medir otro aspecto puntual relacionado con los enfoques como es la percepción que los alumnos tienen del curso que están realizando. Para ello diseñaron varias escalas que incluían objetivos claros y tradicionales, métodos de enseñanza formales, cantidad de trabajo, relevancia vocacional, buena enseñanza, libertad en el aprendizaje, apertura de los estudiantes y clima en el aula. Encontraron que los alumnos con enfoque profundo valoraban más positivamente la buena enseñanza y la libertad en el aprendizaje, mientras que la cantidad de trabajo fue el factor que más se asoció con los alumnos que tenían un enfoque superficial.

Esta percepción por parte de los estudiantes se basó en un consenso entre ellos, pero existen muchos aspectos académicos en los que no se da dicho consenso (por ejemplo, en los sistemas de evaluación). Y existen otros aspectos del ambiente donde se registra claramente una diferencia cualitativa de unos estudiantes a otros. El modo, por ejemplo, de percibir la enseñanza dependerá de la orientación educativa que tenga el alumno (lo que espera del curso). Meyer (1988) ha sugerido que la ausencia de correlación entre la percepción del curso y el enfoque de aprendizaje a nivel individual podría deberse no sólo al tipo de pregunta formulada, sino a las mismas técnicas de análisis correlacional utilizadas. De hecho Meyer y Parsons (1989), Meyer y Muller (1990a y b), Meyer, Parson y Dunne (1990), Entwistle, Meyer y Tait (1991) han demos-

trado que usando los procedimientos del escalamiento multidimensional que denominan 'analysis unfolding' es posible representar en un espacio bi- o tridimensional de forma simultánea el enfoque de aprendizaje y la percepción asociada con el ambiente de aprendizaje.

Entwistle y Tait (1990) describieron la relación que existe entre las preferencias que los estudiantes manifiestan por los tipos de enseñanza, la orientación en el estudio y la preferencia por diferentes ambientes de aprendizaje. Ello explica que una determinada orientación en el estudio vaya acompañada de un enfoque para aprender y una forma concreta de realizar las tareas académicas. Así por ejemplo, los estudiantes que adoptan una orientación hacia el significado usan un enfoque profundo apoyándolo en una motivación intrínseca, manifestando preferencias por un tipo de enseñanza, exámenes, tutoría, tipo de cursos, etc., que en otras investigaciones han demostrado que facilitan el aprendizaje (Entwistle, 1990, Entwistle y Ramsden, 1983). Estas interpretaciones se han considerado coherentes para poder ser interpretadas en términos del modelo de enseñanza-aprendizaje elaborado por Entwistle (1987).

Meyer y colaboradores han introducido una nueva perspectiva en el análisis de este tipo de relaciones sin los supuestos de linealidad implícitos en la técnica de la correlación. El 'unfolding' es una técnica de escalamiento de actitudes en la que las escalas y los individuos se mapifican en un espacio que permite ver qué grupo de dimensiones le atraen más a los sujetos y qué individuos difieren de una forma considerable de la mayoría de su grupo. Cuando la posición de estos estudiantes se representa gráficamente en el espacio creado por el análisis 'unfolding', se observa que está fuera del 'cluster' de la mayoría de los estudiantes, que es normalmente donde se encuentra el área de las escalas. El alejamiento de estos sujetos del resto del grupo y de ellos entre sí se denominan outlier, cuyas características individuales no pueden captarse o representarse en el análisis. Pues bien, al explorar las características de estos estudiantes, Meyer encontró que la mayoría eran sujetos académicamente débiles y el espacio que ocupaban representaba una total desintegración de los modelos esperados en la relación entre enfoques de aprendizaje y la percepción del contexto del aprendizaje. Este nuevo planteamiento ha hecho que Entwistle reanalice algunos datos de sus investigaciones sobre la relación entre orientación en el estudio, la evaluación de la enseñanza y las preferencias por los tipos de enseñanza, los exámenes, la tutoría y el curso.

El resultado más sorprendente al analizar los sujetos propensos al fracaso es que la relación entre el enfoque y las percepciones se hace más interpretable. Esta desintegración de las estrategias de aprendizaje respecto a la percepción del ambiente del aprendizaje parecen ser los hallazgos más novedosos. Es decir, en los alumnos que fracasan se da una incoherencia importante entre su orientación en el estudio y la percepción que tienen de su ambiente de aprendizaje. Biggs (1985) había observado ya la ausencia de una estructura clara de los ítems del SPQ en alumnos con bajos rendimientos.

4. IMPACTO DE LA EVALUACIÓN EN LA CALIDAD DEL APRENDIZAJE

Ya hemos visto en las investigaciones presentadas y en el modelo propuesto por

Biggs que existen una serie de variables que repercuten en la calidad del aprendizaje. Una de tales variables, que aparecen de forma reiterada, es la evaluación. De tal forma es tan importante la evaluación, a juicio de algunos autores, que se ha llegado a generalizar la idea de que los métodos de evaluación son determinantes básicos del tipo de aprendizaje alcanzado por los alumnos. Para Biggs (1996) los procedimientos de evaluación son un determinante del aprendizaje de los alumnos en mayor grado que lo son el currículum y los métodos de enseñanza. Si la evaluación tiene unos efectos tan importantes ¿por qué no mejoramos el aprendizaje de los alumnos simplemente cambiando los sistemas de evaluación? Elton y Laurillard (1979) señalan que el modo más rápido de cambiar la forma en que los alumnos aprenden es cambiando los sistemas de evaluación. Más aún, lo que evaluemos y cómo lo evaluemos demostrará lo que es importante para nosotros como profesores en la formación que esperamos lograr en nuestros alumnos y por consiguiente en los contenidos y procedimientos de evaluación.

Como hemos visto en páginas anteriores, la evaluación del aprendizaje ocupa un lugar destacado en el modelo de enseñanza-aprendizaje propuesto por los partidarios de la perspectiva cualitativa. Este nuevo papel asignado a la evaluación ha producido un corpus de investigación nada despreciable. Autores como Goslin (1967), Kellaghan, Madaus y Airasian (1982), Crooks (1988), etc. han prestado gran atención a este tema.

Crooks (1988) ha llevado a cabo una extensa revisión de las investigaciones sobre el impacto que la evaluación tiene sobre los estudiantes. Sus conclusiones finales apuntan a que la evaluación tiene un impacto directo e indirecto que puede ser positivo o negativo, lo que hace que la evaluación deba planificarse atentamente y cuidar su implementación.

En la revisión de Crooks aparecen cuatro apartados de los cuales dos merecen especial atención. En un primer apartado, el autor revisa la naturaleza, el papel y el impacto que la evaluación tiene en el aula. Según Crooks una gran proporción de tiempo la dedica el alumno a actividades que luego van a ser evaluadas. La gran mayoría de las actividades de evaluación tienen lugar dentro del aula. Los profesores opinan que las actividades de evaluación son aspectos importantes de los procesos de enseñanza-aprendizaje y así lo consideran en su trabajo pero son conscientes de lo inadecuado de su esfuerzo. Una proporción importante de profesores están poco o nada formados en técnicas de evaluación, y aquellos que reconocen poseer dicha formación la encuentran irrelevante para sus actividades evaluativas.

A la pregunta de ¿qué evaluar? muchos profesores están de acuerdo en que hay que evaluar los tres dominios apuntados en la taxonomía de Bloom. La importancia que los profesores asignan a la evaluación de unos dominios sobre otros va a influir en los enfoques de aprendizaje que los alumnos adopten. Por consiguiente, el enfoque de aprendizaje adoptado por el alumno será un predictor importante del resultado de la calidad de dicho aprendizaje. Biggs (1993) ha establecido una relación congruente entre enfoque de aprendizaje (superficial, profundo y de alto rendimiento) y la calidad del aprendizaje resultante (hecho vs estructura). Se establece por tanto una relación interactiva entre evaluación, enfoque de aprendizaje y resultados del aprendizaje:

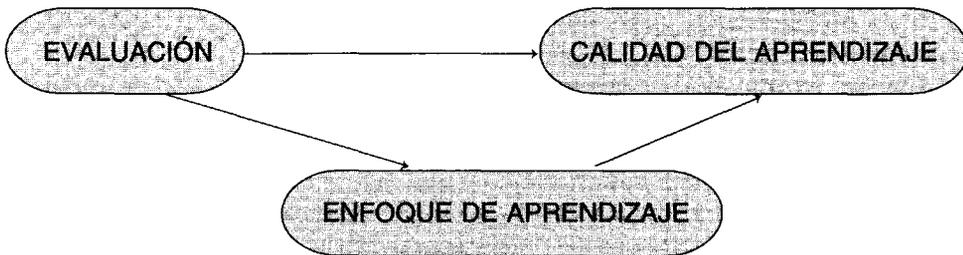


Figura 4

Aún reconociendo los profesores la importancia de incluir en su evaluación los tres dominios de Bloom (aprendizajes cognitivo, afectivos y psicomotores) Fleming y Chambers (1983) consideran que el 80% de las preguntas formuladas por los profesores pertenecen al nivel de los conocimientos. Haertel (1986) encontró también que los exámenes reflejan los objetivos instruccionales que los profesores se han planteado en sus programaciones con un sesgo hacia la repetición de los materiales presentados en clase o en los libros de texto. Crooks observa que las pruebas hechas por los profesores tienden a poner más énfasis en los niveles cognitivos inferiores y, a veces, incluso alejados de los objetivos de la programación.

En la revisión realizada por nosotros sobre un total de 50 programaciones hechas públicas procedentes de diversas titulaciones observamos que la evaluación propuesta por los profesores presenta la misma tendencia señalada por Crooks en su revisión. En la mayoría de los estudios realizados sobre el formato de la evaluación se señala que la única información dada al alumno es sobre si habrá una prueba objetiva o una prueba tipo ensayo. En la revisión que hemos efectuado nosotros obtuvimos el mismo resultado. La información que aparece en las programaciones sólo hace referencia —las que lo hacen— a si se hará una prueba ensayo (con especificación a veces del número de preguntas) o una prueba objetiva (también algunas especifican el número de preguntas). Muy pocos profesores optan por la evaluación oral; y cuando lo hacen queda expresado en términos muy genéricos.

La mayoría optan también por la evaluación sumativa; sólo en algunas programaciones se observa que habrá un seguimiento en lo que son las lecturas y trabajos o sesiones de seminario.

En algunas programaciones se hace una diferenciación, especialmente en las programaciones de las áreas de ciencias, entre la evaluación de la parte teórica y la parte práctica, asignando incluso un peso a cada una de estas partes en la evaluación final.

Prácticamente ninguna programación establece una relación entre los objetivos y la evaluación; de tal forma es así que muy pocos alumnos reciben una guía, al menos de forma escrita en las programaciones, sobre los niveles cognitivos que se van a evaluar en los exámenes. Algunas programaciones incluyen términos como conocimientos adquiridos, comprensión de la materia. El modelo de evaluación utilizado por los profesores afecta sin lugar a dudas al enfoque que los estudiantes adoptan no sólo a

corto sino a medio y largo plazo: en la preparación diaria, en la preparación de las evaluaciones, en la aplicación de lo aprendido, etc.

En síntesis, en este análisis general realizado observamos que sólo en contados casos se dice por escrito al alumno qué se va a evaluar, cómo se le va a evaluar y cuándo se le va a evaluar. Creemos que todos estos aspectos de la evaluación necesitan de un análisis sistemático más profundo y una mayor reflexión.

El segundo apartado lo dedica Crooks a las investigaciones que se han llevado a cabo para estudiar el impacto que las prácticas evaluativas tienen en el aprendizaje y el rendimiento de los alumnos. Las investigaciones parecen apuntar que existe una relación entre los enfoques de aprendizaje y las estrategias de enseñanza utilizadas por el profesor en clase, especialmente los procedimientos evaluativos que utiliza. Marton y Säljö (1987a) fueron los primeros en establecer la categorización de los enfoques de aprendizaje adoptados por los alumnos universitarios y constatar que estos enfoques venían determinados por la forma en que los alumnos eran enseñados y evaluados. A conclusiones similares llegaron otros autores como Entwistle y Ramsden (1983), Laurillard (1984), Ramsden (1985), Elton y Laurillard (1979), etc. La elección del enfoque de aprendizaje parece depender entre otros factores del tipo de evaluación utilizado por el profesor. Elton y Laurillard (1979) llegan a la conclusión de que si queremos cambiar el modo en que los alumnos abordan su aprendizaje y estudio debemos cambiar los sistemas de evaluación. Biggs (1996) ha llegado a una conclusión similar cuando dice que el efecto negativo de la evaluación determina lo que el alumno aprende, mucho más que el currículum y los métodos de enseñanza adoptados por los profesores.

Los efectos de la influencia de la evaluación en el aprendizaje han sido también abordado por Becker et al (1968), Miller y Parlett (1974). Todos han observado que los estudiantes planifican su trabajo con el objetivo último de hacer bien los exámenes y otras tareas evaluativas. Martin y Ramsden (1987), Marton y Saljo (1976b), Ramsden (1984), Van Rossum y Schenk (1984) han observado además que los alumnos que utilizan un enfoque superficial de aprendizaje tienen más dificultades para adaptarse a los requisitos de las evaluaciones que aquellos que utilizan un enfoque profundo. Entwistle y Kozeki (1985) observaron también que la influencia que tiene lo que el profesor recalca en su clase y el modo de evaluarlo influye en el enfoque de aprendizaje de los alumnos.

Todas estas investigaciones han servido para demostrar el gran efecto que la evaluación ejerce en la forma en que el alumno aprende. Ericksen (1983) señala textualmente «un examen es una declaración que revela lo que para un profesor es lo más importante en su materia».

Balla y Boyle (1994) en su artículo «Assessment of Student Performance: a framework for improving practice» señalan que «la evaluación del estudiante y la evaluación del curso y la enseñanza son elementos críticos en el ambiente de la enseñanza-aprendizaje y son centrales en la misión que la institución universitaria tiene en la preparación futura de sus alumnos» (p. 17).

A la vista de las conclusiones que las numerosas investigaciones aportan, parece claro que es responsabilidad de los profesores y de la institución estimular el uso y

desarrollo de estrategias de aprendizaje que lleven a enfoques profundos y de alto rendimiento desde el comienzo de su vida universitaria. La forma en que los profesores diseñen y lleven a cabo la evaluación y la institución la facilite permitirá este tipo de perfiles y como consecuencia una mejor calidad en el aprendizaje.

La evaluación, sin lugar a dudas, ocupa un lugar determinante en la calidad de los resultados del aprendizaje, pero la práctica evaluativa es parte de un esquema más amplio que incluye y va más allá por extensión de la responsabilidad exclusiva de los profesores. Siguiendo a Biggs (1996) entendemos que la institución es un sistema interactivo holístico que para su propia organización dispone de muchos procedimientos, que tiene su propio funcionamiento pero que determina los modelos de enseñanza y de evaluación y que a su vez incide en la percepción que los alumnos tienen de lo que deben aprender y cómo deben aprenderlo. No es suficiente dejar a criterio del profesor que resuelva las demandas conflictivas a nivel burocrático y en lo que a la calidad del aprendizaje se refiere. Cada institución debería tener su propia política y una guía sobre la evaluación suministrando una serie coherente de principios y procedimientos acerca de la evaluación en consonancia con los propósitos y objetivos de la institución y los programas públicos que intentan lograr esos objetivos institucionales.

Uno de los fines de las instituciones universitarias es velar por una formación científica y profesional de sus alumnos y licenciados. Si la evaluación es el dispositivo generador de dicha calidad, creemos honestamente que el Plan Nacional de la Evaluación de la Calidad de las Universidades debería incluir en su dimensión Enseñanza algunas de los resultados y conclusiones presentados más arriba.

BIBLIOGRAFÍA

- ATKINS, J.M., BEATTIE, J. y DOCKRELL, W.B. (1993). *Assessment Issues in Higher Education*. Sheffield: Employment Department.
- BALLA, J. y BOYLE, P. (1994). Assessment of Student Performance: a framework for improving practice. *Assessment and Evaluation in Higher Education*, 19 (1), pp. 17-18.
- BANTA, T.W. (1992). Student achievement and assessment of institutional effectiveness. En Clark, B.R. y Neave, G.R. (eds.) *The Encyclopedia of Higher Education*. Oxford: Pergamon Press.
- BECKER, H.S.; GEER, B. y HUGHES, E.C. (1968). *Making the grade: The academic side of college life*. New York: Wiley.
- BIGGS, J.B. (1993). What do inventories of students' learning processes really measure? A theoretical review and clarification. *British Journal of Educational Psychology*, 63, 3-19.
- BIGGS, J. (1996). Assessing Learning Quality: reconciling institutional, staff and educational demands. *Assessment and evaluation in Higher Education*, 21 (1), pp. 5-15.
- COLE, S.C. (1990). Conceptions of Educational Achievement. *Educational Researcher*, 19(3), pp. 2-6.
- CONSEJO DE UNIVERSIDADES (1994). *Programa experimental de evaluación de la calidad del sistema universitario. Informe final*. Madrid: MEC.

- CONSEJO DE UNIVERSIDADES (1995). *Programa de Evaluación de la Calidad de las Universidades*. Madrid: MEC.
- CONSEJO DE UNIVERSIDADES (1996). *Plan Nacional de Evaluación de la Calidad de las Universidades. Guía de Evaluación*. Madrid: MEC.
- CROOKS, T.J. (1988). The Impact of Classroom Evaluation Practices on Students. *Review of Educational Research*, 58 (4), pp. 438-481.
- DANSEREAU, D.F. (1985). Learning strategy research. In J. Segal, S. Chipman, & R. Glaser (Eds.), *Thinking and learning skills, Vol. 1: Relating instruction to research*. Hillsdale, N. J.: Erlbaum.
- DE MIGUEL, M.; MORA, J.G. y RODRÍGUEZ, S. (1991). *La evaluación de las instituciones universitarias*. Madrid: Consejo de Universidades.
- DE MIGUEL, M. (1991). Utilización de indicadores en la evaluación de la docencia. En De Miguel, M.; Mora, J.G. y Rodríguez, S. (1991) *La evaluación de las instituciones universitarias*. Madrid: Consejo de Universidades.
- DE LA ORDEN, A. y COLB (1986). Un acercamiento experimental a la investigación del rendimiento en la universidad. *Revista de Investigación Educativa*, 8, 21-26.
- ELTON y LAURILLARD (1979). Trends in student learning. *Studies in Higher education*, 4, pp. 87-102.
- ENTWISTLE, N. (1990). Introduction: Changing conceptions of learning and teaching. En Entwistle, N. *Handbook of educational ideas and practice*.
- ENTWISTLE, N. (1992). Student learning and study strategies. En Clark, B. y Neave, G. (eds.) *The Encyclopedia of Higher Education*. Oxford: Pergamon Press.
- ENTWISTLE, N. y RAMSDEN, P. (1983). *Understanding Student Learning*. London: Croom Helm.
- ENTWISTLE, N. y TAIT, H. (1990). Approaches to evaluation, evaluation of teaching, and preference for contrasting academic environments. *Higher education*, 19, 169-94.
- ENTWISTLE, N.; KOZEKI, B. y TAIT, H. (1989). Pupils' perceptions of school and teachers I-Identifying the underlying dimensions. *Br. J. Education Psychol*, 59, 326-339.
- ENTWISTLE, N.; MEYER, J. y TAIT, H. (1991). Student failure: disintegrated patterns of study strategies and perception of the learning environment. *Higher Education*, 21, 249-261.
- FLEMING, M. y CHAMBERS, B. (1983). Teachers-made tests: Windows on the classroom. En W. E. Hathaway (ed.) *New direction for testing and measurement: vol. 19. Testing in the schools*. S. Francisco: Jossey-Bass.
- FRANSMAN, H.J. (1991). Students' preferences for learning from an evaluation point of view. *South African Journal of Higher Education*, vol. 5 (2), pp. 73-81.
- GIBBS, G.; MORGAN, A. and TAYLOR, E. (1984). The world of the learner. In F. Marton et al. (Eds.) *The Experience of Learning*. Edinburgh: Scottish Academic Press.
- GOSLIN, D.A. (1967). *Teachers and testing*. New York: Russell Sage Foundation.
- HAERTEL, E. (1986). *Choosing an using classroom test: Teachers' perspectives on assessment*. Comunicación presentada en el Congreso anual de AERA.
- HAGER, P. y BUTLER, J. (1996). Two Models of educational Assessment. *Assessment and Evaluation in Higher Education*, 21(4), pp. 367-77.

- HERNÁNDEZ PINA, F. (1993). Concepciones en el estudio del aprendizaje de los estudiantes universitarios. *Revista de Investigación Educativa*, 22, pp. 117-150.
- KELLAGHAN, T.; MADAUS, G.F. y AIRASIAN, P.W. (1982). *The effects of standardized testing*. Boston: Kluwer-Nijhoff.
- KNIVETON, B.H. (1996). Student Perception of Assessment Methods. *Assessment and Evaluation in Higher Education*, 21(3), pp. 229-78.
- LAURILLARD, D. (1984). Learning from problem-solving. En Marton, F.; Hounsell, D.J. y Entwistle, N.J. (Eds.) (1984) *The Experience of Learning*. Edinburgh: Scottish Academic Press Oxford Polytechnic.
- MARTON, F. (1981). Phenomenography-Describing conceptions of the world around us. *Instructional Science* 10, 177-200.
- MARTON, F. and SALJO, R. (1976b). On qualitative differences in learning. II Outcome as a function of the learner's conception of the task. *British Journal of Educational Psychology* 46, 115-127.
- MARTON, F. and SVENSSON, L. (1979). Conceptions of research in student learning. *Higher Education* 8, 471-486.
- MARTON, F. and WENESTAM, C.G. (1979). Qualitative differences in the understanding and retention of the main point in some texts based on the principle-example structure. In M.M. Gruneberg; P.E. Morris and R.N. Sykes (Eds.) *Practical Aspects of Memory*. Academic Press.
- MARTON, F.; HOUNSELL, D.J. y ENTWISTLE, N.J. (Eds.) (1984). *The Experience of Learning*. Edinburgh: Scottish Academic Press Oxford Polytechnic.
- MARTON, F. and SÄLJÖ, R. (1976a). On qualitative differences in learning: I. Outcome and process. *British Journal of Educational Psychology*, 46, 4-11.
- MARTON, F. and SÄLJÖ, R. (1976b). On qualitative differences in learning. II Outcome as a function of the learner's conception of the task. *British Journal of Educational Psychology* 46, 115-127.
- MEYER, J. (1983). Academically at risk study behavior: a categorisation procedure and an empirical exploration based on programmes. *Studies in Higher Education*.
- MEYER, J. y MULLER, M. (1990d). An unfolding analysis of the association between perception of learning context and approaches to learning. *South African Journal of Higher Education*, 4, 46-57.
- MEYER, J. y PARSONS, P. (1989). Approaches to studying and course perception using the Lancaster Inventory-a comparative study. *Studies in Higher Education*, 14, 137-135.
- MEYER, J.; PARSON, P. y DUNNE, T. (1990). Individual study orchestrations and their association with learning. *Higher Education*, 20, 67-89.
- MEYER, J. (1988). Academically at risk study behavior: a categorisation procedure and an empirical exploration based on programmes. *Studies in Higher Education*.
- MEYER, J. y MULLER, M. (1990a). Evaluation the quality of student learning. I. An unfolding analysis of the association between perceptions of learning context and approaches to studying at an individual level. *Studies in Higher Education*, 15, 131-153.
- MEYER, J. y MULLER, M. (1990b). Evaluation the quality of student learning. II. Study orchestration and the curriculum. *Studies in Higher Education*, 16, 251-275.

- MEYER, J., PARSON, P. y DUNNE, T. (1990). Individual study orchestrations and their association with learning. *Higher Education*, 20, 67-89.
- MILLER, R.I. (1979). *The Assessment of College Performance*. S. Francisco: Jossey-Bass.
- MILLER, C.M. y PARLETT, M. (1974). *Up to the mark. A study of the examination game*. London: Society for research into Higher Education.
- NUUTINEN, A. (1996). Learning experiences of University student. Draft copy.
- RAMSDEN, P. (1984). The context of learning. En Marton, F.; Hounsell, D.J. y Entwistle, N.J. (Eds.) (1984) *The Experience of Learning*. Edinburgh: Scottish Academic Press Oxford Polytechnic.
- RAMSDEN, P. (1985). Student learning research: retrospective and prospect. *Higher Education Research and Development*, Vol. 4, No 1, 52-69.
- RAMSDEN, P. (1985b). *Alternatives to Learning Skills*. Paper presented at the 6th Australasian Tertiary Study Skills Conference, Adelaide.
- RODRÍGUEZ ESPINAR, S. (1991). Calidad universitaria: un enfoque institucional y multidimensional. En De Miguel, M.; Mora, J.G. y Rodríguez, S. (1991) *La evaluación de las instituciones universitarias*. Madrid: Consejo de Universidades.
- ROSSUM VAN, E.J. y SCHENK, S.M. (1984). The relationship between learning conception, study strategy and learning outcomes. *British J. of Educational Psychology*, 54, 73-83.
- TAYLOR, E. (1984). Orientation to study: A longitudinal investigation of two degrees in one university. Unpublished Ph.D. dissertation, University of Surrey.
- THOMAS, P.R., and BAIN, J.D. (1982). Consistency in learning strategies. *Higher Education*, 11, 249-59.
- VROEIJENSTIJN, T.I. y ACHERMAN, H. (1991). Evaluación de la calidad basada en el control frente a la evaluación de la calidad basada en la mejora. En De Miguel, M. Mora, J.G. y Rodríguez, S. *La evaluación de las instituciones universitarias*. Madrid: Consejo de Universidades.
- WEINSTEIN, C. E. (1978). Teaching cognitive elaboration learning strategies. In H. F. O'Neil, Jr. (Ed.), *Learning strategies*. New York: Academic Press.
- WITTRICK, M. C. (1978). The cognitive movement in instruction. *Educational Psychologist*, 13, 15-29.

LA EVALUACIÓN EN LA EDUCACIÓN SECUNDARIA OBLIGATORIA EN EL NUEVO SISTEMA EDUCATIVO EN ANDALUCÍA

Eugenio Hidalgo Díez
Universidad de Granada

RESUMEN

New Education System is a necessity and a task to teachers because it suppose a reflection of the own planners and practices. In this way, we have treated to analyze the evaluation and the teaching learning model's in new Education System, and also the evaluation of the own teaching-learning process; and finally in order to develop an application model's in tha school with these planners exposed.

INTRODUCCIÓN

Queda ya muy atrás la fecha en la que se comenzó a fraguar la nueva reforma del sistema educativo. A lo largo de este período de tiempo, muchas han sido las opiniones vertidas sobre la misma: favorables, desfavorables, contraproducentes,... y un largo etcétera. No obstante, siempre hemos detectado un denominador común a todos los comentarios: la necesidad de la misma (Buendía, L., 1993), pero a pesar de esta necesidad, por todos admitida, a la hora de poner en marcha todo el proyecto, siempre se veían y se ven dificultades que parecen insalvables.

Hemos de admitir que profesionalmente la implantación del nuevo sistema educativo, además de una necesidad, ya puesta de manifiesto, era y es un reto para todo el profesorado, ya que propugna un cambio en los esquemas y planteamientos didácticos que nos lleven a tratar de conseguir un ciudadano más acorde con la demanda que la sociedad futura necesita.

Esto así dicho resulta complicado, ya que supone para los docentes una reflexión profunda de sus planteamientos y prácticas y una remodelación de ambos de acuerdo con las necesidades y posibilidades del alumno/a y de su entorno.

Igualmente la consideramos un reto, porque nos permite un alto grado de flexibilidad en los planteamientos, así como la novedad de poder organizar nuestro propio diseño curricular para procurar una perfecta adaptación al alumnado que tengamos, teniendo en cuenta las peculiaridades e ideosincrasia propias de cada centro, de su entorno y de los alumnos/as que allí concurren y viven.

Finalmente, indicar, que a lo largo de nuestro planteamiento, hemos partido y estado siempre dentro del marco legislativo ya establecido. De esta manera, tratamos de dar una visión de la evaluación del proceso de enseñanza-aprendizaje en la Educación Secundaria Obligatoria, en el sentido de cómo deberíamos planteárnosla dentro del marco normativo ya establecido y teniendo en cuenta que el proceso de enseñanza-aprendizaje procura desarrollar capacidades por medio de los contenidos específicos de cada área o materia.

I. MARCO TEÓRICO QUE GENERA EL MODELO DE REFORMA

La actual reforma del sistema educativo manifiesta un objetivo explícito: mejorar la calidad de la enseñanza. En coherencia con este objetivo, incluye también una reforma del currículo, que no tiene que entenderse sólo como formal, sino también como una nueva concepción, como un cambio en profundidad, alrededor de unos principios que inspiran y dan lugar a su organización y funcionamiento. Toda práctica educativa ha de estar fundamentada en teorías del aprendizaje y del desarrollo suficientemente válidas, que sirvan como marco teórico desde donde, el profesorado pueda programar, reflexionar y revisar su práctica escolar.

Con respecto a estas teorías, no es nuevo que en un principio se partiese de un enfoque constructivo, quizás porque el devenir del tiempo así lo demandaba y también porque después de una anterior forma de entender la organización del aprendizaje y del conocimiento era necesario una nueva cultura del aprendizaje que diera paso a otra forma de adquisición de dicho aprendizaje y del propio conocimiento y nos sacara del reduccionismo conductista en el que se había caído.

De esta manera, se puede decir que el enfoque constructivo supuso un auge sobre los demás enfoques teóricos y que como tal propugnó un proceso de cambio (Lyddon, 1990; Prochaska y DiClementi, 1986) educativo en la escuela, abogando por un aprendizaje y desarrollo que tuviese en cuenta aquellos aspectos que hacen referencia a dichos procesos, en la dirección de considerarlos como procesuales y constructivos, ya que tanto el aprendizaje como el desarrollo (Kegan, 1982) se construyen progresivamente en interacción con el entorno natural y socio-cultural en combinación con las

posibilidades genéticas del individuo (Piaget, 1970) y el ambiente social en el que se desenvuelve (Erikson, 1959). Por ello, podemos decir, que el aprendizaje supone un proceso de incorporación continua de nuevas informaciones a los esquemas previos que ya se poseen y a la reestructuración de estos, lo que provocaría una reconstrucción continua de los acontecimientos y estos a su vez un proceso de cambio. Como consecuencia, si dichos esquemas no permiten el aprendizaje, se produciría un intercambio con el medio que daría lugar al desarrollo.

Así, el currículum de la L.O.G.S.E., adopta una serie de postulados generales que definen el marco teórico en donde se enmarca, y que no suponen prescripciones de carácter categórico. Estos postulados generales giran alrededor de aspectos como: considerar el desarrollo de la persona como un todo integrado (Piaget, 1975), y que este nivel de desarrollo (Z.D.P.: zona de desarrollo próximo) viene determinado por el aprendizaje y la enseñanza (Vigotsky, 1979), aprendizaje que ha de ser significativo, ya que será el propio alumno quien establezca relaciones entre los conocimientos previos y la nueva información (Ausubel, 1977), para ello, es necesario, que el alumno/a establezca una relación entre el nuevo aprendizaje y el que ya posee, para partiendo de ahí, establecer un desequilibrio entre ambos y crear una conciencia en el alumno/a para superar dicho desequilibrio que favorezca la adquisición del nuevo aprendizaje.

Otro postulado viene determinado por la interacción social, es decir, por una perspectiva interaccionista del desarrollo humano, donde la construcción e interiorización de los contenidos del aprendizaje y su conocimiento se lleva a cabo en un proceso de interacción social (Bruner, 1988; Piaget, 1970; Vigotsky, 1979; Brofenbrenner, 1989).

Con lo que, el currículum del actual sistema educativo genera un nuevo planteamiento del proceso de enseñanza-aprendizaje, pero a su vez, esta nueva forma de concebir la construcción del proceso de enseñanza-aprendizaje se basa en unas concepciones que procuran por todos los medios que el alumno/a «aprenda a aprender», es decir, proveerle de unos mecanismos, para que sea el propio alumno/a el que partiendo de unos conceptos, procedimientos y actitudes, adquiridos previamente sobre el contenido, construya nuevos significados (Ausubel, Novak, Hanesian, 1983) de dicho contenido y a su vez enriquezca y vaya modificando su estructura cognoscitiva, lo que a su vez dará lugar a que sea capaz de desarrollar estrategias que le sirvan para hacer nuevos aprendizajes.

En este sentido, se podría decir que, el currículo de la L.O.G.S.E. adopta explícitamente un determinado modelo psicopedagógico: el constructivismo, ya que: «El alumno es el responsable último de su propio proceso de aprendizaje», igualmente, «La actividad mental constructiva del alumno se aplica a contenidos que poseen ya un grado considerable de elaboración», para finalmente considerar que: «La actividad constructiva del alumno obliga a sustituir la imagen clásica del profesor como transmisor de conocimientos por la del profesor como orientador o guía». (Coll, C., 1990).

Esto no quiere decir que ese aprendizaje se realice sin necesidad de ayuda externa, sino al contrario: la función del profesor se vuelve fundamental, dado que la actividad de autoconstrucción que efectúa el alumno/a se inscribe dentro de una actividad interpersonal. El alumno/a aprende en el contexto de las relaciones con sus compañe-

ros y el profesor, siendo determinante, la actuación de éste para que la actividad del alumno/a sea más o menos constructiva.

No obstante, no se pretende que ante lo expuesto se entienda que la teoría de partida por la que aboga el proceso de reforma sea la de un constructivismo reduccionista y que de esa manera se debiese entender todo el proceso de enseñanza-aprendizaje que se pretende exponer, pues si así fuese se estaría dando lugar a caer en los mismos errores en que se han visto inmersas otras teorías precedentes, sino que por el contrario: «sea lo que sea que acabe pasando con la búsqueda de la verdad, los acontecimientos con los que nos enfrentamos hoy están sujetos a tanta variedad de construcciones como nuestro ingenio nos permita inventar. Esto no quiere decir que cualquier construcción es tan buena como otra... Pero nos recuerda que todas nuestras percepciones presentes están abiertas al cuestionamiento y la reconsideración, y sugiere ampliamente que incluso las ocurrencias diarias podrían aparecer completamente transformadas si tuviéramos imaginación suficiente para construirlas de modo diferente» (Kelly, 1970: 1).

Como se ha apuntado con anterioridad en los diferentes postulados a los que se ha venido haciendo referencia desde un principio, lo que se trata es de procurar aplicar al modelo de enseñanza-aprendizaje que se pretende un sentido constructivo como realmente hay que entenderlo, es decir, no sólo quedándonos: «en un proceso de individuación y construcción de la identidad personal», sino también de: «socialización y de incorporación a una sociedad y una cultura» (Coll, 1996), desde la realidad nuestra de cada día, o sea, desde nuestra familia, microsociedad (barrio, ciudad,...), nuestros intereses, en definitiva desde nuestro entorno que es donde el centro escolar se encuentra inmerso y desde donde se puede empezar a entender, por el contacto que existe o debe existir entre familia/centro, la realidad de lo que cada alumno/a es y puede llegar a ser. De esta manera, hemos de entender el modelo constructivo como una familia interrelacionada de teorías que cuestionan la versión realista y objetivista de la ciencia (Neimeyer, 1996).

En consecuencia, si se ha planteado que en la nueva reforma subyace un marco constructivista en el proceso de enseñanza-aprendizaje, igualmente hemos de entender la evaluación, por lo que, se tendrá en cuenta en el momento de la misma:

- el desarrollo y aprendizaje de las capacidades.
- la detección del grado de significatividad de los aprendizajes.
- el carácter o entidad que le atribuyen los alumnos/as a los aprendizajes.
- la forma en que se produce el proceso de aprendizaje significativo, entendido como un continuum y que por lo tanto la evaluación no debe ser puntual sino producirse en el tiempo que dure el proceso de enseñanza-aprendizaje, y tanto sobre la enseñanza como sobre el aprendizaje al unísono.
- el utilizar el máximo número posible de actividades sobre el contenido para proceder a la evaluación del mismo.
- que el grado de control sobre el aprendizaje vaya pasando del profesor a los alumnos.
- que procure siempre información al alumno/a sobre la construcción de sus aprendizajes.

2. LA EVALUACIÓN Y EL MODELO DE ENSEÑANZA-APRENDIZAJE EN EL ACTUAL SISTEMA EDUCATIVO

La actividad evaluadora está íntimamente ligada a las diferentes maneras de entender el proceso de enseñanza-aprendizaje como apuntábamos anteriormente, por ello, para una mejor comprensión del modelo empezaremos planteando los componentes del diseño curricular y posteriormente el desarrollo del proceso.

a) Componentes del Diseño Curricular

Es necesario para abordar la construcción del proceso de enseñanza-aprendizaje, tener, en principio, una visión general de lo que conocemos como componentes del diseño curricular, de esta manera y en sentido vertical por su interrelación a lo largo de todo el diseño, podemos establecer los siguientes:

- Finalidades Educativas:

La actividad docente que se realiza en un centro educativo, tiene un propósito y busca cómo conseguirlo y es necesario que todo ello quede plasmado como las finalidades educativas del centro. Por ello, consideramos como finalidades educativas, las afirmaciones de principios generales e intenciones educativas acordes con las Finalidades Educativas del Sistema Educativo y adaptadas al propio centro, su entorno y su alumnado.

- Objetivos Generales de la E.S.O.:

Establecidas las Finalidades Educativas y en relación con ellas se encuentran los Objetivos Generales de la E.S.O. que: Precisarán las Capacidades que los alumnos deben haber adquirido al finalizar la etapa correspondiente. En este modelo los objetivos son generales y abiertos; pero esto no significa nada más que la posibilidad de que sean los propios profesores en función de cada situación educativa los que realicen mayores especificaciones, de tal manera que las intenciones educativas pierdan generalidad para ganar concreción en la unidad didáctica.

- Objetivos Generales del Área:

En consonancia con los Objetivos Generales de la E.S.O., obviamente estarán los Objetivos Generales del Área que serán los que: Precisan las capacidades que los alumnos/as deben haber adquirido a través del área o materia correspondiente. La función del objetivo es guiar la práctica y orientar la articulación de los diferentes elementos curriculares en función de las intenciones educativas. Si la guía no es adecuada, no se ajusta a la realidad, los objetivos no cumplirán su función primordial de ser un marco de actuación docente.

Si formulamos el objetivo: Conseguir que los alumnos tengan una actitud positiva hacia los emigrantes y no planteamos marcos de actuación para el desarrollo de capacidades como la tolerancia, adaptación,... etc., puede ocurrir como en el estudio de Berttel (1984) que exponía la siguiente situación: «En una clase de Ética con veinte alumnos blancos y dos negros, de una escuela pública, el profesor pretende conseguir la no discriminación racial y modificar positivamente la actitud hacia los negros. Después de varias sesiones de trabajo en clase, el profesor evaluó la actitud a través de una escala tipo Lickert. Posteriormente clasificó a los alumnos según la puntuación en la escala. Al subirse los alumnos al autobús, de regreso a la casa, a pesar de las buenas puntuaciones en la escala de la mayoría de ellos, los niños negros volvieron a sentarse juntos, y aunque entraron en el autobús en diferente momento, ningún niño blanco ocupó el asiento de al lado del niño negro».

- Criterios de Evaluación:

Por su parte los criterios de evaluación establecen el tipo y grado de aprendizaje que se espera que los alumnos/as hayan alcanzado con respecto a las capacidades implícitas en los Objetivos Generales. Es decir, hacen referencia al desarrollo de las capacidades del alumnado de forma general.

- Capacidades:

Implícitas en los Objetivos tanto Generales de Etapa como de Área, se encuentran las capacidades que nos indican: El potencial o aptitud que posee una persona para llegar a la adquisición de nuevos conocimientos y habilidades. Es decir, es el potencial que se debe desarrollar en los alumnos/as desde los 0 a los 16 años, ya que durante todo el período de escolarización que abarcan estas edades, los grupos de capacidades a desarrollar desde las diferentes áreas y/o materias serán siempre las mismas.

- Contenidos:

El otro aspecto que como docentes debemos procurar en nuestro alumnado es la asimilación de los contenidos, entendiendo estos como: El conjunto de formas culturales y saberes seleccionados y secuenciados para formar parte de las distintas áreas curriculares formulados en base a: conceptos, procedimientos y actitudes.

Los contenidos estarán en función de los objetivos. Un contenido es cualquier elemento de la realidad que contribuya a desarrollar las capacidades expresadas en los objetivos tanto Generales de Etapa como los específicos de cada área. El desarrollo de estas capacidades demandará la enseñanza de diferentes tipos de contenidos (conceptuales, procedimentales y actitudinales) y consecuentemente diferentes formas e instrumentos de evaluación. El conocimiento del contenido

ELEMENTOS BÁSICOS DEL CURRÍCULO

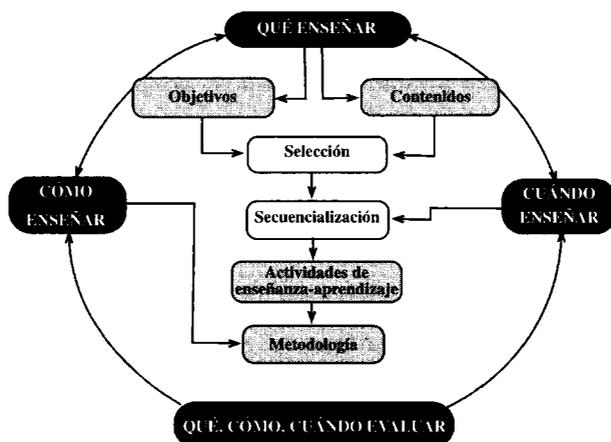


Gráfico 1

Elementos Básicos del Currículo.
 (Grupo Asesores CEP-Granada, 1993).

de las materias, como único conocimiento de interés para ser evaluado, no es más que una forma de justificar intereses docentes que en absoluto se corresponden con los intereses del alumno ni para el alumno.

Los elementos claves de la evaluación sobre los que reflexionaremos es este trabajo son: Cuándo evaluar, para qué evaluar y cómo evaluar.

- **Objetivos Didácticos:**

En nuestra manera de construir el proceso de enseñanza-aprendizaje, este componente es el fundamental, ya que lo consideramos como: El indicador del tipo y grado de aprendizaje que debe realizar el alumnado y el referente del proceso de evaluación de los aprendizajes.

b) Construcción del proceso de enseñanza-aprendizaje

Una vez que se han analizado los componentes principales del Diseño Curricular del Área, es decir del proceso de enseñanza-aprendizaje, y antes de plantear el modelo de construcción, es importante realizar las siguientes reflexiones:

- a) La construcción del proceso de enseñanza-aprendizaje debe ser acordada y consensuada por el conjunto del profesorado, ya que implica la toma de decisiones sobre una serie de elementos, tales como: qué, cuándo y cómo enseñar y qué, cuándo y cómo evaluar que son esenciales para el desarrollo de la actividad docente.

- b) Para la elaboración del proceso de enseñanza-aprendizaje, también se ha de tener en cuenta una serie de referentes, tales como: Los Decretos, ya que establecen el currículum de la etapa, especificando los Objetivos Generales de la Etapa y los Objetivos Generales del Área, así como otras normativas. Como podemos deducir, todos ellos son puntos de referencia, en la medida que constituyen el referente normativo en el actual sistema.

Otro referente a tener en cuenta sería el conjunto de documentos y materiales que poseen los profesores/as (programaciones, documentos, planificaciones previas,... etc.) que hayan sido realizados en otro momento y que merecen una consideración positiva del profesor.

Para la construcción del proceso de enseñanza-aprendizaje, se han establecido dos fases. En la realización de la primera fase del proceso, se pueden establecer los siguientes momentos:

1. Analizar y establecer la relación entre los Objetivos Generales de Etapa y las Capacidades, es decir, indicar qué capacidad o capacidades desarrollan los diferentes objetivos.
2. Analizar y establecer la relación entre los Objetivos Generales de Etapa y los Objetivos Generales de Área y/o Materia.
3. Como consecuencia de las dos relaciones anteriores, podemos establecer, directamente, que capacidad o capacidades se desarrollan en los diferentes Objetivos Generales de Área.

De esta manera al estar relacionados los Objetivos Generales del Área con los Objetivos Generales de Etapa y las Capacidades, sabremos qué capacidades estamos trabajando en cada momento de forma automática, debido a la interrelación establecida en el primer momento de la construcción del proceso de enseñanza-aprendizaje.

Hacemos esto, porque en la propuesta curricular establecida en el Diseño Curricular Base de la Consejería de Educación y Ciencia de la Junta de Andalucía, se ha optado por expresar los resultados de la intervención educativa en términos de capacidades y no de conductas observables (D.C.B. pág. 38).

Además, porque a la hora de la evaluación del proceso, tendremos que valorar no sólo el grado de asimilación de los contenidos, sino también el tipo y grado de desarrollo de las capacidades.

La segunda fase se centra en la realización de una Programación de Aula o Unidad Didáctica, con la ventaja de no haber perdido en ningún momento, debido a la interrelación que debe procurarse, los referentes fundamentales del proceso de enseñanza-aprendizaje, como son los Objetivos Generales de Área en los que, como ya hemos dicho, están implícitas las Capacidades.

De esta manera, en primer lugar, se establece la relación entre los Objetivos Generales de Área y los Contenidos en sus tres aspectos: Conceptuales, Procedimentales y Actitudinales.

Una vez establecida esta relación, el siguiente paso trataría de aunar de alguna

forma las capacidades implícitas en los Objetivos Generales de Área y los Contenidos que ya previamente hemos relacionado.

Esto se consigue con la formulación de lo que denominamos Objetivos Didácticos. Este es el momento más importante y clave en la construcción del proceso de enseñanza-aprendizaje. Porque son:

- Los que aunan Capacidades a desarrollar y Contenidos a asimilar.
- El referente fundamental para la construcción de la Unidad Didáctica.
- Los que marcan la pauta para la organización de las actividades con nuestros alumnos/as.
- Los que pueden ser objeto de evaluación y por consiguiente, hacemos ver el grado de desarrollo de capacidades y a la vez, la asimilación del contenido por parte del alumnado.

Es decir, que los Objetivos Didácticos dan forma y entidad a la Unidad Didáctica, ya que en ellos tenemos el referente de Capacidades y Contenidos. Si partiendo de esto, organizamos para dichos Objetivos Didácticos, unas Actividades, las Temporizamos en sesiones, diseñamos una Metodología para su aplicación en clase y Evaluamos dichas actividades, habremos construido la Unidad Didáctica o Programación de Aula, habiendo aunado los dos referentes de evaluación:

- Grado de desarrollo de Capacidades.
- Grado de asimilación de Contenidos.

3. LA EVALUACIÓN DEL PROCESO DE ENSEÑANZA-APRENDIZAJE

Establecido el modelo de enseñanza-aprendizaje entraremos en el análisis de la evaluación, teniendo en cuenta:

- a) El concepto implícito en el marco legal.
- b) La evaluación como elemento curricular.
- c) Intenciones y objetivos que pretende alcanzar.
- d) Aspectos del proceso de enseñanza-aprendizaje afectados por la evaluación.

a) Concepto implícito en el marco legal

La evaluación es una actividad valorativa e investigadora y, por ello, facilitadora de cambio educativo y desarrollo profesional docente afecta a:

- Los procesos de aprendizaje del alumnado.
- Los procesos de enseñanza del profesorado.
- Al Proyecto Curricular de Centro.

Constituyendo de este modo, el elemento clave para:

- Orientar las decisiones curriculares.
- Definir los problemas educativos.
- Acometer actuaciones concretas.
- Empezar procesos de investigación educativa.
- Generar dinámicas de formación permanente del profesorado.
- Regular, en definitiva, el proceso de adaptación y contextualización del currículum en cada comunidad educativa.

Por lo que, la evaluación, desde este marco, se define como:

* Integradora.

* Continua.

* Formativa.

* Cualitativa.

* Contextualizada.

Integradora, ya que exige tener en cuenta las capacidades generales establecidas en los Objetivos de la Etapa, a través de las expresadas en los Objetivos de las distintas Áreas o Materias.

Continua, al ser un elemento inseparable de todo el proceso de enseñanza-aprendizaje que aporta información del mismo con carácter permanente. Por tanto estará presente de forma sistemática, en el desarrollo de todo tipo de actividades y no sólo en momentos puntuales y aislados.

Formativa, pues al tener información de los procesos de enseñanza-aprendizaje, con carácter permanente, puede detectar las dificultades en el momento que se producen, averiguar las causas y en consecuencia regular, orientar y autocorregir la intervención educativa con objeto de mejorar tanto los procesos seguidos por todos los participantes.

Cualitativa, al ofrecer datos e interpretaciones sobre los cuales poder entender y valorar los procesos seguidos por todos los participantes.

Contextualizada, ya que tiene en cuenta el contexto del Centro, su Proyecto Curricular, así como la singularidad de cada uno de los alumnos, considerando su propio proceso de aprendizaje, sus características y sus unidades específicas.

Desde este modelo podemos decir que evaluar es obtener información acerca de como se está desarrollando el proceso de enseñanza-aprendizaje con objeto de ayudar al alumno, a adaptarse a sus características y necesidades, comprobar la acción didáctica del docente y determinar hasta que punto se han desarrollado las intenciones educativas expresadas en cada proyecto curricular.

b) La Evaluación como elemento curricular

Se entiende por currículum el conjunto de objetivos, contenidos, métodos pedagógicos y criterios de evaluación de cada uno de los niveles, etapas, ciclos, grados y modalidades del sistema educativo que regulan la práctica docente (LOGSE, art. 4.1), es decir, que el currículum tiene dos funciones fundamentales: Hacer explícitas las

intenciones del sistema educativo y servir como guía para orientar la práctica pedagógica. Este «carácter continuado, procesual, contextual y estratégico de la evaluación reguladora en el proceso de enseñanza-aprendizaje es especialmente útil para poner de relieve los aspectos intencionales, deliberativos, reflexivos y de toma de decisiones característicos del uso de procedimientos y el grado de dominio conseguidos por el alumno» (Gómez, I., 1996: 61). Así pues, la evaluación se convierte en un elemento fundamental e inseparable de la práctica educativa que permite recoger la información de cada momento y realizar los juicios de valor necesarios para la orientación y toma de decisiones respecto al proceso de enseñanza-aprendizaje.

c) Intenciones y objetivos que pretende alcanzar la evaluación («El por qué y el para qué evaluar»)

Los procesos del aprendizaje del alumnado, los procesos de enseñanza del profesorado, es decir, la acción educativa, los Proyectos Curriculares de Centro, etapa, ciclo y/o aula así como el Sistema Educativo en general, precisan de una valoración que ayude a su comprensión y análisis para realizar una toma de decisiones que permita mejorar su funcionamiento.

Esta finalidad última se proyecta en una serie de intenciones que la evaluación pretende alcanzar y que entre otras, podemos señalar:

- Referidas a los procesos de aprendizaje:
 - Obtener información y valorar situaciones, procesos y resultados.
 - Detectar dificultades.
 - Tomar decisiones.
 - Reorientar procesos de enseñanza-aprendizaje.

- Centradas en el alumnado:
 - Conocer el grado inicial de desarrollo de las Capacidades.
 - Conocer la situación de partida en los aprendizajes.
 - Conocer la evolución del proceso de enseñanza-aprendizaje.
 - Atender a las necesidades individuales del alumnado.
 - Apreciar y valorar el grado de desarrollo alcanzado en cuanto a las capacidades y contenidos enunciados en los objetivos generales de la etapa y/o ciclo.
 - Organizar las posibles adaptaciones curriculares ante las dificultades detectadas, y en su defecto participar en el proceso de diversificación curricular.
 - Determinar la titulación y promoción del alumnado.
 - Garantizar la movilidad del alumnado de un centro a otro.

- Referidas a los procesos de enseñanza:
 - Obtener información y valorar la actuación docente.
 - Analizar las estrategias metodológicas realizadas durante el proceso de enseñanza con objeto de contribuir al perfeccionamiento de las actuaciones docentes y modificar su práctica.

- Referidas al Proyecto Curricular:
 - Recogida de información y valoración sobre el desarrollo del proyecto Curricular.
 - Toma de decisiones y reorientación sobre el Proyecto Curricular con objeto de mejorarlo y adecuarlo al entorno y contexto de centro y a las características del alumnado.
 - Facilitar a las familias la información necesaria sobre el proceso de aprendizaje de sus hijos/as y la orientación para la continuidad de su formación.
- Referidas al propio Sistema Educativo:
 - Establecer mecanismos de retroalimentación y apoyos necesarios para la mejora del proceso de enseñanza-aprendizaje.

Además de estas intenciones, la evaluación procura una serie de objetivos fundamentales tales como:

- Contribuir a la mejora de la actividad educativa.
- Señalar el grado en que se van desarrollando las diferentes capacidades y orientar las medidas de refuerzo o adaptaciones curriculares necesarias.
- Señalar el grado de asimilación de los contenidos que configuran las diferentes áreas y/o materias.
- Modificar aquellos aspectos de la práctica docente y del Proyecto Curricular que se hayan detectado como poco adecuados a las características del alumnado y al contexto del Centro.
- Ser un instrumento al servicio del proceso de enseñanza-aprendizaje, integrado en el quehacer diario del aula y del centro educativo.
- Ser el punto de referencia para adoptar decisiones que afecten a la intervención educativa y a la mejora del proceso.

d) Aspectos del proceso de enseñanza-aprendizaje afectados por la evaluación. («El qué evaluar»)

Desde el concepto de evaluación anterior decíamos que esta afecta fundamentalmente a:

- Los procesos de aprendizaje del alumnado. En este sentido, hemos de:
 - Evaluar los conceptos previos. Antes de iniciar cualquier aprendizaje hemos de conocer la situación actual o de partida en la que se encuentra el alumnado frente al mismo. Es necesario partir de los conceptos, representaciones y conocimientos previos que ha construido el alumno/a en el transcurso de sus experiencias ya que éstos le sirven como punto de partida y como medio para interpretar la nueva información.
 - Evaluar el grado de desarrollo de las Capacidades y la asimilación de los Contenidos. El objetivo de la evaluación es valorar las capacidades expresadas en

los Objetivos Generales tanto de Etapa como de Área, utilizando como medio para su consecución la asimilación de los contenidos de dicha área.

Antes de continuar, es necesario, debido a la importancia que tiene en el contexto de la evaluación, analizar: ¿Cómo evaluar las Capacidades? Según el nuevo sistema educativo podemos hacerlo de forma indirecta teniendo en cuenta dos tipos de indicadores:

- Los Criterios de Evaluación.
- Los Objetivos Didácticos.

— Criterios de Evaluación:

La evaluación del aprendizaje de los alumnos se realizará teniendo en cuenta los objetivos de etapa, así como los objetivos, contenidos y criterios de evaluación de las distintas áreas del currículo.

Dada la dificultad de evaluar los objetivos, expresados en términos de capacidades, los criterios de evaluación se convierten en un indicador más concreto de la misma.

Los criterios de evaluación proporcionan una información sobre los aspectos a considerar para determinar el tipo y grado de aprendizaje que haya alcanzado el alumnado, en cada uno de los momentos del proceso en cuanto al avance en la adquisición de las capacidades, con respecto a las que ya poseía el propio alumno/a y establecidas en el currículo como capacidades básicas de cada área. Dichos criterios no son directa ni unívocamente evaluables. Es necesario planificar algunas tareas, lo más variadas posibles, para comprobar que las capacidades y contenidos expuestos en los criterios se han alcanzado en el grado propuesto. No se pretende desde los criterios evaluar todo lo que el alumnado aprende, sino que en ellos se seleccionan las capacidades básicas de aprendizaje desarrolladas y sin las que un alumno/a difícilmente avanzaría de forma satisfactoria en su normal proceso de aprendizaje.

Los Equipos Docentes deberán adecuar los criterios de evaluación a los objetivos que establezcan en su Proyecto Curricular de Centro. Estos criterios deben secuenciarse por ciclos de tal modo que las valoraciones acerca de los procesos de aprendizaje se realicen de forma continua no permitiendo que los alumnos/as acumulen retrasos en el desarrollo de las capacidades básicas de aprendizaje.

— Objetivos Didácticos:

Son considerados como: El indicador del tipo y grado de aprendizaje, referido este, tanto a la asimilación de los contenidos como al desarrollo de las capacidades en relación con dichos contenidos de la unidad didáctica correspondiente. Los Objetivos Didácticos guiarán la intervención educativa, constituirán el punto de referencia inmediato de la evaluación continua y permitirán encontrar los procedimientos de evaluación más adecuados a tales objetivos.

Los Objetivos Didácticos se establecerán en cada Unidad Didáctica o Programa-

ción de Aula. por su grado de concreción evalúan todos los aprendizajes programados. Será a través de los objetivos didácticos y de las actividades programadas para alcanzarlos, desde donde se obtenga la información para el seguimiento y evaluación de todos los contenidos.

La evaluación de los procesos de aprendizaje se realiza a través del seguimiento de las actividades diseñadas para el logro de los Objetivos Didácticos en las Unidades Didácticas o Programaciones de Aula, por lo que los Objetivos Didácticos serán el referente más cercano para la valoración del día a día de la actividad del aula.

- Los procesos de enseñanza del profesorado:

En este sentido hemos de:

- * Evaluar el proceso de enseñanza y la práctica docente a través de los siguientes aspectos:
 - La organización del centro y el aprovechamiento de los recursos.
 - El carácter de las relaciones y de la convivencia entre profesor y alumnos/as.
 - La coordinación entre los órganos y personas responsables en el Centro de la planificación y desarrollo de la práctica docente.
 - La coordinación entre el profesorado que compone el Equipo Docente, para establecer el tipo y grado de desarrollo de los diferentes grupos de capacidades en los Objetivos Generales de Etapa.
 - La coordinación entre el profesorado que componen los diferentes Departamentos de Área, para establecer el tipo y grado de desarrollo de las capacidades y a través de que contenidos se van a realizar.
 - La estructuración de las adaptaciones curriculares.
 - La regularidad y calidad de la relación con las familias.

- El Proyecto Curricular:

En este sentido hemos de tener en cuenta los siguientes aspectos:

- Adecuación de los Objetivos Generales de Etapa a las Finalidades del Centro y estas a su vez a las necesidades y a las características del alumnado y de su entorno.
- Correcta organización del proceso de enseñanza-aprendizaje en el sentido expuesto.
- Adecuación de la Metodología y de los materiales didácticos.
- Establecimiento adecuado de los grupos de capacidades, de su desarrollo y de los criterios de evaluación.
- Validez de la secuenciación de contenidos y de las estrategias de evaluación.
- Adecuación de las medidas de adaptación y diversificación curricular, así como de la orientación educativa y profesional.

4. APLICACIÓN DEL MODELO EN EL AULA

Siguiendo lo ya expuesto para la construcción del proceso de enseñanza-aprendizaje, queremos hacer las siguientes consideraciones, al objeto de que podamos comprobar que dicho modelo ha procurado acercarse lo más posible a estos planteamientos, no con el ánimo de que sea algo definitivo y conclusivo, sino como un punto de reflexión para establecer opciones, que abran paso a diferentes planteamientos constructivos en educación con respecto al proceso de enseñanza-aprendizaje en la escuela, en el sentido de que: «La finalidad de la concepción constructivista no es explicar el desarrollo y el aprendizaje humano. Su finalidad es distinta; su finalidad es configurar un esquema de conjunto orientado a analizar, explicar y comprender los procesos escolares de enseñanza y aprendizaje» (Coll, 1996: 176).

De esta manera, cuando se planteó la construcción del proceso de enseñanza-aprendizaje, establecimos dos fases, como podemos apreciar en el siguiente gráfico:

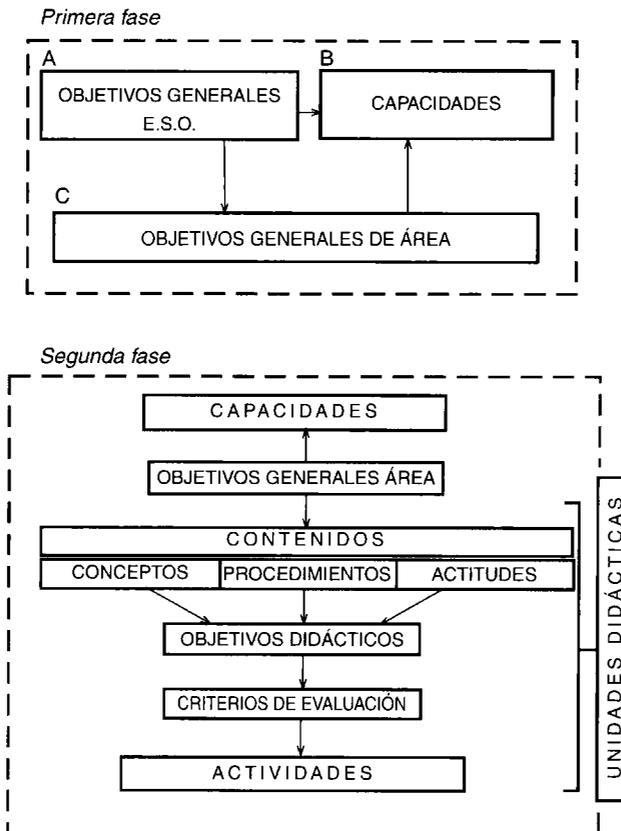


Gráfico 2

Construcción del proceso de enseñanza-aprendizaje.
(Hidalgo, E. y Almagro, J.L.).

No obstante, antes de comenzar el proceso de construcción, es fundamental la definición de las Finalidades educativas del Centro, pues nos van a proporcionar una definición exacta de las características del mismo y de su entorno en sus diferentes ámbitos: Pedagógico, convivencial y de gestión, para lo cual, previamente se han elaborado unos indicadores que nos van a permitir analizar aspectos propios del centro y de su entorno, de los profesores, padres y alumnos.

Una vez que el centro se ha definido, es importante establecer los Objetivos Generales de Etapa, posiblemente, haya que redefinir algunos o todos los objetivos propuestos en el Decreto de E.S.O., pues es fundamental que nuestros objetivos estén adaptados a las Finalidades del Centro, ya que sino estaríamos construyendo un proceso de enseñanza-aprendizaje irreal y no adaptado a los alumnos/as y al entorno en el que viven.

Realizada la definición del Centro, el proceso sería el que ya hemos explicitado en los apartados anteriores, tan sólo destacar, al objeto de dar sentido a esta forma de construcción del proceso los siguientes aspectos.

- Al definir el centro, para establecer sus finalidades, se empieza un proceso de interacción entre los diferentes componentes de la comunidad escolar, tanto los directamente implicados (padres, profesores y alumnos) como los más indirectos (entorno, asociaciones,... etc.), que den lugar a que existan «unos valores propios de la escuela, reflejo de su identidad y propósitos, que son compartidos por sus componentes» (OCDE, 1991).
- Al establecer las diferentes Capacidades a desarrollar:
 - a. Se procurará que estas estén de acuerdo con la realidad del alumnado que el centro tiene y será, en este momento, cuando se empiece o se deba empezar a compensar las desigualdades que traen nuestros alumnos/as por razones del entorno socio-cultural-económico, tanto familiar como del propio barrio, en el que viven, con ello, se quiere decir, que no podemos establecer el mismo tipo y grado de desarrollo de capacidades para los alumnos/as de un centro que se encuentra ubicado en una zona residencial de una gran ciudad, que en un centro de la misma ciudad ubicado en una zona suburbial-periférica de características marginales. Posiblemente, en el primer centro, siempre hablando con carácter general, el alumnado del mismo, tendrá, con respecto a capacidades de interacción, integración y atracción social, o a capacidades de desarrollo afectivo, equilibrio personal y autoestima, o a capacidades morales o éticas un grado de desarrollo, que presumiblemente, ha de ser mucho mayor que el del alumnado del segundo centro al que se ha hecho referencia, por ello, quizás se deba en este centro hacer más hincapié en el desarrollo de estas capacidades, sin que esto quiera decir que debamos abandonar el de las demás.
 - b. Es necesario establecer un trabajo en equipo por parte del profesorado, para determinar y priorizar que Capacidades hay que potenciar desde las diferentes áreas y/o materias, redefinir los Objetivos Generales de la Etapa, si fuese necesario, y en definitiva dar coherencia a todo el proceso de enseñanza-aprendizaje.

OBJETIVOS GEN. ETAPA	CAPACIDADES	OBJETIVOS GEN. ÁREA	CONTENIDOS	OBJETIVOS DIDÁCTICOS	ACTIVIDADES	EVALUACIÓN
<ul style="list-style-type: none"> • Conocer y comprender los aspectos básicos del funcionamiento del propio cuerpo y la incidencia que tienen diversos actos y decisiones personales tanto en la salud individual como en la colectiva. 	<p>I. Para el desarrollo cognitivo o intelectual.</p> <p>II. Para el desarrollo corporal y campo de la salud.</p>	<ul style="list-style-type: none"> • Conocer y valorar su cuerpo y contribuir a mejorar sus cualidades físicas básicas y sus posibilidades de coordinación y control. 	<p>CONCEPTOS.</p> <ul style="list-style-type: none"> • Conocimiento de los conceptos de condición física y sus componentes. • Factores que inciden sobre la condición física: edad, estado físico, capacidades heredadas. • Relación entre los conceptos de salud física y condición física. <p>PROCEDIMIENTOS.</p> <ul style="list-style-type: none"> • Realización de actividades físicas individuales y colectivas, con diferentes niveles de intensidad. <p>ACTITUDES.</p> <ul style="list-style-type: none"> • Toma de conciencia de la propia condición física y responsabilidad en el desarrollo de la misma. • Valoración de los efectos positivos que sobre la salud psicofísica tienen los trabajos de condición física. 	<p>El alumno/a es capaz de identificar el concepto de condición física y sus componentes.</p> <p>El alumno/a es capaz de analizar los factores que inciden sobre la condición física.</p> <p>El alumno/a es capaz de relacionar los conceptos de salud física y condición física.</p> <p>El alumno/a es capaz de realizar actividades físicas individuales y colectivas, con diferentes niveles de intensidad.</p> <p>El alumno/a es capaz de apreciar su propia condición física y ser consciente del desarrollo de la misma.</p> <p>El alumno/a es capaz de percatarse de los efectos positivos que tienen los trabajos de condición física sobre la salud.</p>	<p>A VER SI ERES CAPAZ. (Ver página siguiente).</p> <p>ANALIZA Y ANALIZATE. (Ver página siguiente).</p> <p>RELACIONA CONDICIÓN FÍSICA Y SALUD. (Ver página siguiente).</p>	
<p style="text-align: center;">CRITERIOS DE EVALUACIÓN</p> <ul style="list-style-type: none"> • Haber logrado una evolución cualitativa (mejora de la flexibilidad, precisión) de las cualidades físicas básicas con respecto a uno mismo. • Haber conseguido ser consciente de lo beneficioso que es para la salud física un adecuado desarrollo de la condición física. • Ser consciente de nuestras posibilidades y limitaciones en la práctica de la actividad física, tanto individuales como colectivas. 						

ACTIVIDADES (I)	ACTIVIDADES (II)
<p>A VER SI ERES CAPAZ: Entre las siguientes definiciones que te propongo, identifica la que corresponde a la de Condición Física y rodeala con un círculo azul.</p> <p>a. Conjunto de cualidades anatómicas y fisiológicas y motoras que reúne una persona para poder realizar esfuerzos físicos diversos.</p> <p>b. La capacidad que permite realizar movimientos de gran amplitud.</p> <p>c. Conjunto de cualidades anatómicas que reúne una persona, para poder realizar esfuerzos físicos diversos.</p> <p>ANALIZA Y ANALÍZATE: * Normalmente, en una prueba de fuerza quien levantará más peso un hombre de 70 años u otro de 25 años. ¿Por qué? * Un atleta bien preparado de 24 años o una persona de la misma edad. ¿Por qué? * Dos personas de la misma edad. ¿Por qué? Analiza las respuestas que has dado y contesta: ¿Qué factores deduces de las contestaciones que inciden e influyen sobre la Condición Física?</p> <p>RELACIONA CONDICIÓN FÍSICA Y SALUD: Rodea con una línea roja aquellas oraciones que creas que hacen referencia a la Condición Física y de azul las que se refieran a Salud Física.</p> <p>* Ingerir una gran cantidad de alimentos. * Sentir la necesidad de hacer actividad física. * No practicar ningún tipo de actividad física. * Desarrollar a través de la actividad física todas mis cualidades anatómicas, físicas y motoras. * Darse cuenta de que cuando hago alguna actividad física mi estado, en general mejora.</p> <p>COMPRUEBA TU FUERZA: — Tumbado en el suelo mirando hacia arriba y con las piernas extendidas:</p> <ul style="list-style-type: none"> • Eleva las piernas hasta la vertical y bájalas. • Eleva, ahora, el tronco hasta la misma posición. <p>— En ambos casos la musculatura que han trabajado es la misma.</p> <p>a). ¿Cuál de los dos movimientos te cuesta más? b). ¿Por qué?</p>	<p>COMPRUEBA TU CAPACIDAD AERÓBICA:</p> <ul style="list-style-type: none"> • Intenta recorrer en 12 minutos la máxima distancia que puedas. No es necesario que cambies de ritmo, y puedes caminar si te encuentras cansado: pero es preferible que regules tus fuerzas para hacer el recorrido de una manera más o menos regular. • Anota tus pulsaciones al finalizar. • Compara la distancia que has recorrido con las distancias de otras personas de tus mismas características. • ¿Sabes que existe una relación entre la frecuencia cardíaca y la resistencia? Recuerda que: * Hasta 140 p/m. Resistencia Aeróbica. * Entre 140 p/m y 170 p/m. Resistencia Aeróbica-Anaeróbica. * A partir de 170 p/m. Resistencia Anaeróbica. <p>REFLEXIONA Y VALORA: Contesta sí o no: Antes de hacer tus prácticas de resistencia: — ¿Te planteabas que tipo de resistencia ibas a trabajar? — ¿Te tomabas las pulsaciones antes de la práctica? Después de realizar tus prácticas de resistencia: — ¿Te planteas que tipo de resistencia vas a trabajar? — ¿Te tomas las pulsaciones antes de la prueba?</p> <p>RESPONDE: — La salud es la capacidad para responder positivamente a las circunstancias del ambiente: deporte, vida sana, alimentación sana y hábitos saludables son los factores de buena salud. — Cuando sales con los amigos, los fines de semana, seguro que habrás observado que la gente practica hábitos no muy encaminados a mejorar su salud: describe algunos comportamientos que creas que no son buenos para la salud...</p>

- También se establece la secuenciación de los contenidos y su relación con las Capacidades implícitas en los objetivos.
- Se ha de tener en cuenta que en los Objetivos Didácticos se agruparán las Capacidades y los Contenidos, al objeto de que al organizar las actividades que desarrollen dichos Objetivos Didácticos, podamos saber que capacidades estamos desarrollando y a través de que contenidos lo estamos haciendo.

De forma resumida, podemos apreciar lo expuesto de forma práctica a través de un pequeño ejemplo, en la tabla 1.

En este sentido, las actividades que se organizan para desarrollar los Objetivos Didácticos correspondientes deberán hacer referencia a los correspondientes indicadores, tanto de aprendizaje, como de desarrollo de capacidades, como de asimilación del contenido, que nos permitan, por medio de la estrategia que nos hayamos propuesto (Observación de la clase, corrección de tareas,... etc.), efectuar una evaluación continua del proceso.

Igualmente se tendrá en cuenta, para llevar a cabo lo expuesto con anterioridad, que cuando se organicen los Objetivos Didácticos, se tienen que hacer constar los indicadores esperados de aprendizaje, que a su vez, son los que, al mismo tiempo, nos van a indicar que capacidad queremos desarrollar y a continuación que contenido queremos que se asimile.

Así, si decimos que:

«El alumno/a es capaz de identificar el concepto de fuerza». (Objetivo Didáctico).

Y organizamos una actividad, por ejemplo:

«De entre las definiciones que te propongo identifica, rodeándola de amarillo, la que hace referencia a la definición de fuerza». (Actividad de clase).

Ante el resultado se podrá inferir que:

- Si el alumno/a ha sido capaz de identificar el concepto correcto de fuerza, de entre otros varios erróneos, es que dicho alumno/a tiene aprendido dicho concepto (aprendizaje).
- Si el alumno/a ha sido capaz de identificar correctamente, es que desarrolla adecuadamente capacidades del tipo cognitivo-intelectual (desarrollo).
- Si el alumno/a ha identificado correctamente el concepto de fuerza, es que dicho contenido lo tiene asumido (conocimiento).

Igualmente será fácil darse cuenta, que si en el desarrollo de actividades, se puede inferir los aspectos que se han apuntado, como aprendizaje, desarrollo y asimilación de los contenidos, esto mismo será lo que apreciemos cuando realicemos el proceso de evaluación de dichas actividades.

Finalmente, se puede apreciar que en la construcción del proceso de enseñanza-aprendizaje que proponemos no se quiere quedar en un constructivismo reduccionista, sino que se intenta «incorporar y utilizar instrumentos conceptuales y metodológicos que tienen su origen en otros campos disciplinares» (Coll, 1996: 177), por ejemplo, cuando tratamos de definir nuestro centro y analizar las reglas y el orden social en sus aulas; cuando se habla de pautas de interacción social; cuando se habla de desarrollo de capacidades desde las diferentes áreas y/o materias; cuando «los profesores trabajan en equipo, colaboran en la planificación, coparticipan en la toma de decisiones»; «cuando el currículo se planifica cuidadosamente e incluye tanto las materias que permitan a los alumnos adquirir los conocimientos y destrezas básicos como las indicaciones para una evaluación continuada, al tiempo que refleja los valores asumidos por la escuela» (OCDE, 1991).

De esta manera, y dentro del marco del nuevo sistema educativo, se ha pretendido, por un lado, dar respuesta a la normativa vigente, por otro lado, intentar que el profesorado pueda desarrollar estrategias para llevar a cabo un proceso que de alguna manera le empiece a implicar en la filosofía del nuevo sistema educativo y que respondiese a su principal objetivo, es decir, mejorar la calidad de la enseñanza, en el sentido de: «Planificar, proporcionar y evaluar el currículum óptimo para cada alumno, en el contexto de una diversidad de individuos que aprenden» (Wilson, 1992: 34).

También se ha pretendido, por último, abrir una línea a través de la cual se puedan aunar la teoría y la práctica, así como la de un enfoque más abierto y multidisciplinar y procurar la participación de los docentes y la interacción de todos los agentes sociales que intervienen en el proceso de construcción de los aprendizajes, desarrollo y conocimientos del alumno/a.

BIBLIOGRAFÍA

- ANTÚNEZ, S.; DEL CARMEN, L.M.; IMBERNÓN, F.; PARCERÍA, A. y ZABALA, A. (1992). *Del proyecto educativo a la programación de aula*. Graó. Barcelona.
- AUSUBEL, NOVAK y HANESIAN (1983). *Psicología educativa, un punto de vista cognitivo*. Trillas. México.
- BLANCO PRIETO, F. (1994). *La Evaluación en la E.S.O.* Amarú. Salamanca.
- BROFENBRENNER, V. (1989). *La ecología del desarrollo humano*. Paidós. Barcelona.
- BRUNER, J.S. (1988). *Desarrollo cognitivo y educación*. Morata. Madrid.
- BUENDÍA EISMAN, L. (1993). *Análisis de la Investigación Educativa*. Universidad de Granada. Granada.
- CENTRO DE PROFESORES (1993). *Evaluación del proceso de enseñanza-aprendizaje y del proyecto curricular: Instrumentos*. C.E.P. Granada.
- COLL, C. (1987). *Psicología y Currículum*. Paidós. Barcelona.
- COLL, C.; PALACIOS, J. y MARCHESI, A. (1990). *Desarrollo psicológico y educación II*. Alianza. Madrid.
- COLL, C. (1996). Constructivismo y educación escolar. *Anuario de Psicología*, 69, 153-168.

- CONSEJERÍA EDUCACIÓN Y CIENCIA. JUNTA DE ANDALUCÍA (1992): *Decreto de Educación Secundaria*. Junta de Andalucía. Sevilla.
- ERIKSON, E. (1959): *Identity and the life cycle*. International University Press. Nueva York.
- GÓMEZ, I. (1996). Enseñanza y aprendizaje. *Cuadernos de Pedagogía*, 250, 54-64.
- HIDALGO, E. y ALMAGRO, J.L. *Construcción y evaluación del proceso de enseñanza-aprendizaje en Educación Secundaria Obligatoria*. (En edición).
- KEGAN, R. (1982): *The evolving self: Problem and process in human development*. Harvard University Press.
- KELLY, G.A. (1970). A brief introduction to personal construct theory, en D. Bannister (comps.): *Peerspectives in personal construct theory*. Acadec Press. Nueva York.
- LORENZO VICENTE, J.A. (1992). *La Educación Secundaria*. Alhambra Longman. Madrid.
- LYDDON, W.T. (1990). First and second order change: Implications for rationalist and constructivist cognitive therapies. *Journal of Counseling and Development*, 69, 122-127.
- M.E.C. (1993). *Documentos de apoyo a la evaluación en E.S.O.* Servicio de publicaciones del M.E.C. Madrid.
- NEIMEYER, G.J. (1996): *Evaluación Constructivista*. Paidós. Barcelona.
- OCDE (1991). *Cómo valorar la calidad de la enseñanza*. MEC. Madrid.
- PIAGET, J. (1970). *La epistemología genética*. A. Redondo. Barcelona.
- PROCHASKA, J.O. y DICLEMENTI, C.C. (1986). *The transtheoretical approach*. En J.C. Norcross. Brunner/Mazells. Nueva York. (pp. 163-200).
- ROTGER, B. (1990). *Evaluación formativa*. Cincel. Madrid.
- VYGOTSKI, L.S. (1979): *El desarrollo de los procesos psicológicos superiores*. Crítica. Barcelona.

LA EVALUACIÓN DEL PROFESORADO. UN TEMA A DEBATE

Joan Mateo (Universidad de Barcelona)
Tomás Escudero (Universidad de Zaragoza)
Mario de Miguel (Universidad de Oviedo)
José Ginés Mora (Universidad de Valencia)
Sebastián Rodríguez (Universidad de Barcelona)

RESUMEN

La evaluación del profesorado es una de las áreas de la investigación evaluativa con más tradición y a la que se ha dedicado y se dedican más esfuerzos, pero al mismo tiempo es también una de las más conflictivas, sobre todo cuando se proyecta en el terreno de la toma de decisiones y en el de la puesta en marcha de programas de evaluación.

A través de nuestro artículo planteamos en toda su complejidad, la problemática general en torno al tema y sus posibles respuestas, así tratamos aspectos tales como: La caracterización de la función docente como contenido clave a evaluar, la definición de los criterios de calidad de los profesores, los propósitos y consecuencias del proceso evaluador, sin olvidar el reflexionar en torno a las dificultades de aceptación de modelos evaluativos que no ofrezcan claros beneficios potenciales para los profesores y para su trabajo.

Hemos tratado de presentar la evaluación del profesorado vertebrada a través del eje formado por la propia evaluación, los procesos de formación permanente-innovación y el desarrollo profesional.

Finalmente se aborda en el artículo la evaluación de la función directiva y el planteamiento institucional respecto de los procesos de acreditación de directores de centros educativos.

SUMMARY

The assessment of the teaching staff is one of the areas of assessment research with more tradition and one into which more effort has been put, whilst being, at the same time, one of the most problematic, especially when referring to the field of decision-making and the putting into operation of assessment programmes.

In our article we set out, in all their complexity, the general problems surrounding the subject together with possible solutions, dealing with aspects such as the nature of the duties of the teaching staff as a key aspect to be assessed, the definition of the criteria of quality in respect of teachers, the aims and consequences of the assessment process without forgetting to take into consideration the difficulties in getting assessment models accepted when they do not offer clear potential advantages for teachers and their work.

We have endeavoured to present the assessment of the teaching staff by dividing in into the three areas which together constitute the crux of teacher assessment self-assessment, continuing training processes-permanent, innovation and professional development.

Finally the article touches upon the assessment fo the role of management and the official position regarding the processes of recognition of the heads of educational centres.

INTRODUCCIÓN

A través del artículo que vamos a presentar, os proponemos un recorrido que iniciaremos con la descripción de las dificultades que presenta la operativización de cualquier modelo de evaluación del profesorado. Analizaremos posteriormente algunas de las respuestas que se han propuesto en torno a los contenidos, los propósitos, los aspectos metodológicos y culturales que arman y condicionan el tema. Seguiremos con la exposición de la normativa y los principios de intervención, estos últimos básicamente concentrados en los procesos de formación continuada. Reflexionaremos respecto de las vías de acción (o de contención) que se nos antojan como más efectivas y adecuadas y trataremos finalmente de establecer los parámetros básicos sobre los que consideramos viable, la configuración e implantación de un modelo evaluativo del profesorado.

Cabe añadir, que parte de la información utilizada, proviene de un estudio financiado por el CIDE, finalizado en el año 1995 (Mateo y otros), llevado a cabo por el mismo equipo de autores de este artículo.

CONTEXTUALIZACIÓN DE LA ACCIÓN EVALUADORA. DIFICULTADES DE PARTIDA

El profesorado tomado individualmente y muy especialmente como colectivo, constituye uno de los ejes principales sobre el que gravita la calidad de cualquier sistema educativo y su evaluación se está convirtiendo en tema capital en las sociedades occidentales.

Se produce, sin embargo, un dilema fundamental que condiciona el diseño de

cualquier praxis evaluativa, por un lado se considera la evaluación del profesorado, tanto por parte de los expertos como de la comunidad educativa, como el elemento clave en la mejora del sistema educativo y al tiempo, a nadie se le escapa que su formalización es tremendamente compleja y problemática, por no decir que se presume conflictiva, no encontrándose soluciones totalmente satisfactorias entre las ensayadas por gran parte de los países de nuestro contexto socio-cultural más cercano.

En nuestro país la LOGSE preve su evaluación en el artículo 62, inscrito en el título cuarto de la ley, llamado «de la calidad de la enseñanza». Parece, por tanto, que la opción tomada trata de fundamentar o justificar la acción evaluativa ligada directamente a una posible mejora de la calidad de la enseñanza.

Sin embargo el término «calidad de la enseñanza» no está definido operativamente; posiblemente en el futuro, el Instituto Nacional de Calidad y Evaluación lo concrete, pero no es ningún secreto que detrás del mismo subyace un constructo muy complejo, en cuya configuración intervienen numerosos agentes y entre los que se dan una ingente multiplicidad de interacciones.

Surge de momento, el temor que desde la administración, en un exceso de simplificación (por no decir de simplismo), se identifique calidad educativa con rendimiento escolar y se quiera responsabilizar en exclusiva al profesorado de la cadena posible de petición de responsabilidades en torno al funcionamiento global del sistema educativo, ignorando por un lado, la amplitud conceptual que abarca el término y por otro, la inexactitud de atribuir los resultados académicos a los docentes, olvidando que se hallan determinados también, por la acción de otros factores tales como: «procedencia social de los alumnos, condiciones de los centros, funcionamiento de los mismos y por la propia política educativa» (Gimeno, 1993, p. 24).

Por otra parte, es evidente que la ordenación del sistema educativo comporta, en el seno de un estado de derecho, la obligatoriedad de la administración de ejercer su control. Y aquí es donde, a nuestro entender, se origina otra gran área problemática, la que surge de la necesidad de institucionalizar la acción evaluadora con un objetivo sustancialmente distinto al anterior, en esta ocasión para ejercer el debido control sobre el sistema en cada uno de sus apartados y en el que el del profesorado se considera uno esencial.

En este sentido, se confunde, incluso por parte de los propios especialistas, la naturaleza de la acción controladora respecto de la evaluadora. Rul (1993) distingue con gran acierto entre los conceptos de «control-regulación» y «evaluación». «El control-regulación es responsabilidad exclusiva del poder legítimamente constituido y su misión es posibilitar el mantenimiento de la organización a través de la equilibración entre la norma (expectativa social) y la realización individual; la evaluación incide en el aprendizaje social de valores (interiorización), en la conformación cultural, su puesta en marcha corresponde a las fuerza sociales» (Rul, 1993, p. 20).

La dificultar de generar, formalizar y operativizar un modelo evaluativo está, por tanto, en la capacidad de lograr el conveniente equilibrio entre los dos aspectos (control y evaluación) y establecer un sistema relacional justo y eficiente entre ambos.

Dificultad que se complica por la ya mencionada distinción entre la naturaleza de los dos integrantes del modelo. Así la acción evaluadora es crítica y supone siempre ofrecer

visiones no simplificadoras de las realidades evaluadas; posibilita la interpretación de los hechos, el diagnóstico real y en profundidad de los problemas y es orientadora de la intervención. Bien aplicada debe ser siempre generadora de cultura evaluativa.

El control exige previamente a su activación, un importante grado de simplificación, no porque sea simplista, sino porque opera sobre la realidad instrumental. Su acción si es correcta genera cultura normativa, complemento necesario a la evaluativa para facilitar y hacer efectiva la acción mejoradora de la realidad educativa y vertebrar el cambio en profundidad.

Varias son, por tanto, las previsible dificultades con las que hay que contar, cuando se estructure en la práctica un plan evaluador del profesorado (Mateo, J., 1993, Mateo, J. y otros (1995), Escudero, T. (1996):

a) *De orden conceptual*, centradas fundamentalmente en la dificultad de determinar los referentes que permitan substantivar la evaluación y los criterios evaluativos a aplicar.

b) *De orden técnico-metodológico*, Las diversas técnicas metodológicas, desde la autoevaluación hasta los tests de capacidad o competencia pasando por las medidas indirectas, observaciones o entrevistas, no acaban de convencer al profesorado.

c) *De cultura evaluativa*, ¿Cómo se establece una cultura de la evaluación entre el profesorado? Tal como señala Harris (1986), el reto de un programa de evaluación no se ubica en el terreno meramente técnico. La evaluación no es simplemente un argumento procesual, es sobre todo un instrumento generador de cultura evaluativa (Mateo 1995).

d) *De gestión política de la evaluación*, ¿Dónde se inserta la evaluación del profesor dentro del contexto general de la acción educativa en un sistema escolar determinado? Con frecuencia, los procesos de evaluación se asimilan a la simple recogida de información.

e) *De legitimación de la evaluación*, ¿Cuál es el marco legal idóneo dentro del que legitimar los propósitos, el alcance y las consecuencias que se derivan del proceso evaluador y en el que establecer las garantías necesarias para preservar los derechos de los evaluados?

f) *De orden deontológico*, ¿Cómo preservar y garantizar el derecho a la intimidad y al honor de los evaluados a la par que ejercer el debido control y las acciones que de él se deriven tal, como legítimamente le corresponde a la sociedad?

CONTENIDOS SUSTANTIVADORES DE LA EVALUACIÓN DEL PROFESORADO

a) La función docente

La puesta en marcha de un proceso evaluador exige en primer lugar, focalizar adecuadamente la dimensión o dimensiones que mejor configuran el objeto o sujeto a evaluar. En nuestro caso, resulta obvio que no tratamos de evaluar al profesor como persona, sino como profesional docente.

Tal como señalábamos anteriormente, consideramos injusto derivar la evaluación del profesorado desde su responsabilización casi en exclusiva de la calidad de la enseñanza (el análisis de la misma debería realizarse desde el marco general de la evaluación del sistema educativo y no de la del profesorado). Su aportación es singular y significativa pero indirecta, si de algo cabe responsabilizarle es, a nuestro juicio, de la calidad del desarrollo de las funciones que le caracterizan como un buen profesional de la docencia.

Pero... ¿cuáles son los contenidos definidores de la función docente? La literatura al uso ofrece numerosos listados de los aspectos que saturan el término y que deberían constituir los indicadores claves en cualquier evaluación de profesores. Muchos de ellos son reiterativos y para evitar repeticiones nos limitaremos a exponer únicamente dos, cuyas presentaciones se nos antojan como complementarias:

Así y tomando como punto de referencia el «rol» profesional de los profesores, Scriven (1988 y 1994) (en Mateo, J. y otros 1995 y Escudero T. 1996) lleva a cabo un análisis detallado y propone los siguientes conocimientos y competencias básicas como los que debe reunir un buen profesor:

1. Conocimiento de la materia:

- a) en las áreas de su competencia.
- b) en materias transversales del currículum.

2. Competencias instruccionales:

- a) destrezas comunicativas.
- b) destrezas de gestión:
 - gestión del proceso.
 - gestión del progreso.
 - gestión de las emergencias.
- c) destrezas de programación y desarrollo:
 - planificación del curso.
 - selección y creación de materiales.
 - utilización de los recursos disponibles.
 - evaluación del curso, de la enseñanza, de los materiales y del currículum.

3. Competencias de evaluación:

- a) conocimientos sobre evaluación de los estudiantes.
- b) destrezas de construcción y administración de pruebas.
- c) clasificación/puntuación/calificación:
 - proceso.
 - resultado.
- d) registro e informe del rendimiento de los estudiantes:
 - conocimiento sobre informes de rendimiento.
 - proceso de información a estudiantes, administradores, padres y otros.

4. Profesionalidad:

- a) ética.
- b) actitud.
- c) mejora.
- d) servicio:
 - conocimiento acerca de la profesión.
 - ayuda a los profesores noveles y colegas.
 - trabajo para las organizaciones profesionales.
 - investigación sobre la enseñanza.
- e) conocimiento de los deberes.
- f) conocimiento de la escuela y su contexto.

5. Otros deberes con la escuela y la comunidad.

Por su parte Schalock y otros (1993) (en Mateo, J. y otros 1995 y Escudero, T.(1996)), ofrece una perspectiva algo distinta, partiendo del análisis de la productividad y la efectividad docente:

- 1) Conocimiento docente: Preparación en aquellas materias relacionadas con su profesión. Materia a enseñar, pedagogía, psicología del desarrollo, etc.
- 2) Destrezas docentes: Medios que los profesores usan para aplicar su preparación para la docencia; métodos y técnicas didácticas, estímulos, etc.
- 3) Competencia docente: Preparación específica para cumplir de manera adecuada con las responsabilidades y las funciones profesionales tales como diseño y planificación docente, motivación de los estudiantes, orientación y control del progreso de los estudiantes, trabajos con los colegas, administradores y padres.
- 4) Efectividad docente: La contribución del profesor a las ganancias de los estudiantes dentro de un área concreta de contenidos, en un cierto periodo de tiempo.
- 5) Productividad docente: La contribución del profesor a las ganancias de los estudiantes en todo el ámbito formativo, en un cierto periodo de tiempo.
- 6) Profesionalidad docente: Seguimiento de una conducta propia de un verdadero profesional de la docencia: cumplimiento profesional, trabajo con colegas, directivos, padres, servicios a la comunidad escolar, integración en actividades de perfeccionamiento profesional, autocrítica profesional, etc.

De forma muy sintética resumiríamos de manera más comprensiva lo anterior, agrupando los contenidos propios de la función docente, en tres grandes áreas:

- a) *Docencia*
- b) *Profesionalidad*
- c) *Servicios a la Comunidad.*

Bajo el primer epígrafe se sitúan a su vez, tres grandes subáreas: el dominio sobre

las áreas de conocimiento a impartir, la cualidad didáctica empleada en la tarea docente y el uso adecuado de los conocimientos de orden psico-pedagógico.

Dentro del segundo gran apartado encontramos toda una gama de actitudes por parte del profesor que pasan por el cumplimiento escrupuloso de sus obligaciones docentes, el cuidado de su formación continua a través de tareas de formación, reflexión crítica, investigación e innovación y el respeto a los fundamentos y actitudes éticas en el ejercicio profesional.

Finalmente cada vez más se entiende que la tarea docente debe trascender a lo puramente escolar, el profesor debe ser un miembro activo en el marco de su comunidad socio-educativa, por ello es habitual el ver situado por parte de numerosos estudios esta tercera dimensión (servicio a la comunidad) como fundamental en la definición de la función docente.

Sin embargo en la cultura real del actual profesorado, el peso o importancia que se concede a cada uno de los contenidos anteriores, es decididamente desigual. Así se desprende al menos del estudio ya mencionado (Mateo y otros 1995), en el que se encuestaron a más de 1.800 profesores de secundaria, pertenecientes a seis comunidades autónomas (Aragón, Asturias, Cantabria, Cataluña, Comunidad Valenciana y País Vasco).

En él, ante la cuestión de cuáles son los contenidos que mejor definen la función docente, la categoría más escogida fue la docencia en clase, seguida del cumplimiento de las obligaciones propias (más bien hacen referencia a las rutinarias) del profesor. Aparecen en segundo lugar las tareas de perfeccionamiento e investigación-innovación. Se muestran finalmente muy críticos a incluir, el desempeño de funciones y cargos y los servicios a la comunidad, como elementos constitutivos del constructo.

b) La función directiva

Por otro lado la mejora de la calidad de la enseñanza exige ampliar los límites de la evaluación, para que pueda ser aplicada de modo efectivo al conjunto del sistema educativo y es en la actualidad un hecho aceptado que con la evaluación de la función docente del profesorado no agotamos las posibilidades evaluativas del personal docente.

En los sistemas educativos desarrollados se admite, cada vez más, la importancia y la relevancia de la función directiva y su incidencia en el desarrollo y mejora del sistema escolar en particular y del educativo en general y de aquí la creciente atención que se concede a los procesos de selección, formación y evaluación de los equipos directivos de los centros escolares.

Existe una amplia literatura en torno a las funciones del director, en general se entiende que éstas estarán siempre condicionadas al modelo y a la cultura organizativa donde se insertan. Conviene en consecuencia que el directivo analice y reflexione previamente en torno a las características de la organización en la que actúa, a fin de incidir positivamente en la estructura, en el funcionamiento, en la cultura, en las interrelaciones y en las interacciones de dicho contexto escolar.

En síntesis y de una forma muy descriptiva, podríamos agrupar las funciones directivas en las siguientes (Gairin y Darder, 1994):

Funciones de *planificación* general del centro, que en el caso de nuestro país se concreta en la elaboración del Plan Anual.

Funciones de *organización* que implican fundamentalmente la definición y distribución de tareas, y la asignación de responsabilidades entre los miembros de la comunidad escolar.

Funciones de *coordinación*, armonización y sincronización de la actividad del conjunto de personas con los planes y la asignación de recursos económicos.

Funciones *ejecutivas*, tomar decisiones, dar instrucciones, realizar trabajos, aplicar recursos y solucionar problemas.

Funciones de *control y evaluación* de la eficacia del proceso educativo del centro.

De las funciones anteriores se derivan las competencias que le corresponden al director, y que en el marco de la LOPEG (Ley orgánica para la evaluación y el gobierno de los centros docentes), son claramente establecidas en su artículo 21 y que citamos a continuación:

1) Dirigir y coordinar todas las actividades del centro hacia la consecución del proyecto educativo del mismo, de acuerdo con las disposiciones vigentes y sin perjuicio de las competencias atribuidas al Consejo Escolar del centro y a su Claustro de Profesores.

2) Ostentar la representación del centro y representar a la Administración educativa en el centro, sin perjuicio de las atribuciones de las demás autoridades educativas.

3) Cumplir y hacer cumplir las leyes y demás disposiciones vigentes.

4) Colaborar con los órganos de la Administración educativa en todo lo relativo al logro de los objetivos educativos del centro.

5) Designar al Jefe de Estudios, al Secretario, así como a cualquier otro órgano unipersonal de gobierno que pueda formar parte del equipo directivo, salvo al administrador, y proponer sus nombramientos y ceses a la Administración educativa competente.

6) Ejercer la jefatura de todo el personal adscrito al centro.

7) Favorecer la convivencia en el centro e imponer las correcciones que correspondan, de acuerdo con lo establecido por las Administraciones educativas y en cumplimiento de los criterios fijados por el Consejo Escolar del centro.

8) Convocar y presidir los actos académicos y las reuniones de todos los órganos colegiados del centro y ejecutar los acuerdos adoptados en el ámbito de su competencia.

9) Autorizar los gastos de acuerdo con el presupuesto del centro, ordenar los pagos y visar las certificaciones y documentos oficiales del centro.

10) Realizar las contrataciones de obras, servicios y suministros de acuerdo con lo que se establece en la Ley.

11) Cuantas otras competencias se le atribuyan en los correspondientes reglamentos orgánicos.

Podríamos sintetizar lo anterior señalando que independientemente del enfoque más burocrático y detallado marcado por el artículo 21 de la LOPEG, coincidimos con las funciones señaladas por Darder y Gairin, aunque encontramos a faltar alguna dimensión que añadimos a continuación, dejándolas concretadas en las siguientes áreas de actividad:

Planificación.
Organización.
Coordinación.
Ejecución.
Evaluación.
Control.
Información.
Animación.

EL PROPÓSITO EVALUADOR

El proceso evaluador precisa ser orientado por el propósito que lo impulsa y lo justifica, en el caso específico de la evaluación del profesorado se reconocen los siguientes:

La mejora de la calidad de la enseñanza.
El desarrollo profesional del docente.
El reconocimiento social y económico de su labor.
El control y la consiguiente petición de responsabilidades.
La selección de profesores.

Es evidente que no todos los elementos mencionados son de la misma importancia y que están totalmente interrelacionados. Cualquier práctica evaluadora acostumbra a incorporar varios de ellos y deberá tener muy en cuenta sus interacciones más significativas.

Así entendemos que la justificación última de cualquier proceso evaluador, desde la perspectiva social, está en la mejora de la calidad de la enseñanza, pero ésta pasa necesariamente entre otros ámbitos, por el desarrollo profesional del docente tanto en su dimensión individual como colectiva.

No hay desarrollo profesional si no se produce simultáneamente un crecimiento personal y de la cultura grupal del colectivo institucional al que se pertenece.

Por otro lado el desarrollo del profesor va generalmente ligado a incentivos que pasan por el desarrollo de la propia carrera profesional y pretenden entre otras cosas, el reconocimiento laboral, social y económico de su labor.

Finalmente es obvio que la evaluación aporta elementos informativos respecto del docente que pueden ser utilizados por la comunidad para su selección o para el control del ejercicio de su función. Los mismos elementos son indispensables en los procesos de autorreflexión y retroalimentación profesional del docente.

Sin embargo, de los diferentes propósitos mencionados y teniendo presente sus interrelaciones, entendemos que en situaciones en la que no hay una tradición consolidada (ni tan siquiera incipiente) no es prudente iniciar los procesos evaluativos direccionados desde todos sus posibles propósitos, a nuestro juicio consideramos como el más nuclear y prioritario el que plantea la evaluación ligada al *desarrollo profesional*.

Orientar los procesos evaluativos desde su perspectiva más formativa abre un sinfín de posibilidades para racionalizar los procesos de mejora del profesor en tanto que profesional y consecuentemente mejorar la calidad de la enseñanza, finalidad última considerada como la más universal.

En la actualidad las administraciones educativas de los países mejor situados respecto del tema, están claramente evolucionando desde la evaluación ligada fundamentalmente al control y a la mejora salarial (pago por méritos o resultados) a la evaluación entendida como un proceso unido al del perfeccionamiento individual e institucional del profesorado y sustancialmente asociado a su desarrollo profesional.

Se configura, en definitiva, la evaluación de la función docente y de la función directiva como uno de los elementos caracterizadores del desarrollo profesional, conectado a la mejora de las expectativas profesionales y al reconocimiento social como base de los incentivos profesionales y sirviéndose de los procesos de perfeccionamiento personal e institucional como los instrumentos básicos de intervención.

Según Riegle (1987) en Mateo J. y otros (1995) y Escudero T. (1996), este desarrollo puede concretarse alrededor de cinco grandes áreas:

1. *Desarrollo instruccional* que enfatiza las destrezas técnico-pedagógicas.
2. *Desarrollo profesional* que enfatiza la mejora personal dentro de su tarea profesional.
3. *Desarrollo de la organización* que enfatiza las necesidades y prioridades institucionales.
4. *Desarrollo de la carrera* que enfatiza la preparación para el ascenso profesional.
5. *Desarrollo personal* que enfatiza la mejora individual en diversas facetas académicas, vitales, relaciones personales, etc.

Cualquier modelo evaluativo del profesorado que pretenda institucionalizarse, deberá tener en cuenta el cubrir las cinco áreas que concretan el propósito genérico enunciado como «desarrollo profesional». Descuidar alguno de estos aspectos provocará inevitablemente desequilibrios y disfunciones en la estructura de cualquier praxis evaluadora que queramos implantar.

NORMAS EN LA EVALUACIÓN DEL PROFESORADO

El Joint Committee mantuvo para la evaluación del profesorado las mismas normas que las ya propuestas en 1981 para la evaluación de programas, es decir las normas de *legitimidad, utilidad, factibilidad y precisión*.

Las normas de legitimidad hacen referencia a los derechos de las personas afectadas por la evaluación, comprometiendo a que ésta se lleve a cabo de acuerdo a

principios marcados por la ética y la legalidad, actuando siempre en beneficio de los evaluados:

a) **Orientación hacia el servicio:** la evaluación de los profesores debe estar al servicio de los principios educativos, de los mínimos institucionales, del rendimiento en la acción educativa y del logro de las necesidades educativas de los estudiantes y de la sociedad.

b) **Seguimiento de obligaciones formales:** se deben acordar y establecer de manera clara los procedimientos bajo los que se lleva a cabo la evaluación.

c) **Manejo correcto del conflicto de intereses:** se deben identificar los conflictos de interés y tratar de manera abierta y honesta, de manera que no comprometan los procesos y resultados de la evaluación.

d) **Uso limitado de los informes de evaluación:** debe asegurarse el uso apropiado de la información, limitándola a los individuos con una legítima necesidad de usarla y revisarla.

e) **Interacción con los evaluados:** se debe tratar a los evaluados de manera correcta y profesional, de manera que se refuerce su autoestima, motivación, rendimiento, etc. y no se perjudique su actitud hacia la evaluación.

Las de utilidad persiguen que las evaluaciones tengan incidencia en la mejora de la actividad profesional de los docentes:

a) **Orientación constructiva:** debe ayudar a las instituciones y sus profesores a mejorar sus servicios.

b) **Usos definidos:** se deben identificar los objetivos de la evaluación para planear las cuestiones adecuadas.

c) **Credibilidad del evaluador:** el proceso debe ser dirigido y ejecutado por personas calificadas profesionalmente de forma que sus informes sean respetados y creídos.

d) **Funcionalidad de los informes:** deben ser claros y precisos, de forma que sean valiosos para los evaluados y otras audiencias interesadas.

e) **Seguimiento e impacto:** debe existir un proceso de seguimiento para que los usuarios y los evaluados puedan recibir ayudas para comprender los resultados y actuar en consecuencia.

Las normas de factibilidad proponen la necesidad de una evaluación adecuada, eficaz, eficiente y realista, fácil de utilizar y viable social y políticamente:

a) **Procedimientos prácticos:** los procedimientos deben producir la información necesaria, minimizando molestias y costes.

b) **Viabilidad política:** el sistema de evaluación deberá desarrollarse y dirigirse con la colaboración constructiva de todas las partes implicadas.

c) **Viabilidad fiscal:** se debe proporcionar tiempo y recursos a las actividades de evaluación para que se desarrollen de manera efectiva y eficiente,

Finalmente las normas de precisión exigen que la información evaluativa sea relevante, válida y fiable:

a) Definición del contenido a evaluar: las responsabilidades, los objetivos, las cualificaciones, etc. de los evaluados deben definirse claramente, para que el evaluador pueda determinar criterios válidos de evaluación.

b) Análisis del contexto: la evaluación se realiza en un contexto que condiciona el rendimiento evaluado; aquel debe identificarse y describirse.

c) Descripción de los procedimientos: los procedimientos utilizados deben documentarse para que los evaluados y otras audiencias puedan valorarlos a la vista de los propósitos.

d) Medidas válidas: ajustadas al contenido a evaluar y a su potencial utilización, de manera que las inferencias sobre el evaluado sean válidas y precisas.

e) Medidas fiables: los procedimientos de medición deben ser consistentes y fiables.

f) Control sistemático de los datos: la información debe ser bien guardada, cuidadosamente procesada y mantenida.

g) Control de los sesgos: se debe asegurar la protección contra toda información afectada por cualquier tipo de parcialidad.

h) Control del sistema de evaluación: los procedimientos deben revisarse periódica y sistemáticamente.

En el marco legal del Estado Español, la evaluación del profesorado queda regulada mediante el título III de la Ley orgánica 9/95 (LOPEG) y que en los tres apartados de su artículo 30 especifica lo siguiente:

1. A fin de mejorar la calidad educativa y el trabajo de los profesores, las Administraciones educativas elaborarán planes para la valoración de la función pública docente.

2. En la valoración de la función pública docente deberán colaborar con los servicios de inspección los órganos unipersonales de gobierno de los centros y, en los aspectos que específicamente se establezcan, podrán colaborar los miembros de la comunidad educativa que determine la Administración correspondiente. En todo caso, se garantizará en este proceso la participación de los profesores.

3. El plan finalmente adoptado por cada Administración educativa deberá incluir los fines y criterios precisos de la valoración y la influencia de los resultados obtenidos en las perspectivas profesionales de los profesores de los centros docentes públicos. Dicho plan deberá ser conocido previamente por los profesores.

PROCEDIMIENTOS, FUENTES Y AGENTES

Cualquier modelo evaluativo del profesorado deberá decidir respecto de una serie de elementos configuradores de tipo estructural y metodológico. A la elección de los contenidos a evaluar, a la selección de los propósitos últimos que justifiquen y orienten todo el proceso y a la elaboración del marco jurídico que legalice el proceso

evaluador, habrá que añadir una amplia variedad de otros ingredientes metodológicos como son los procedimientos, las fuentes y los agentes evaluativos, que deberán ser introducidos en el modelo y deberán ser acordes con los contenidos y con los propósitos del mismo. (Denham, 1987).

En cuanto a los procedimientos o estrategias de recogida de información evaluativa, técnicamente son aceptables todos los que ofrece la metodología actual, no obstante aparecen como más utilizados y hasta nos atreveríamos a decir más idóneos: los cuestionarios, las entrevistas, la observación y los informes hétéro y autoevaluativos.

Respecto de las fuentes, parece razonable la posición de los profesores expresada en el estudio anteriormente mencionado (Mateo y otros 1995), en la que prefieren aquellas provenientes de su entorno más próximo (propio profesor, compañeros y alumnos) que a nuestro entender, deberán ser complementada con análisis de los documentos de trabajo del profesor y observaciones sistemáticas llevadas a cabo por equipos de evaluadores mixtos (internos y externos). En todo caso las informaciones deberían ser cruzadas, las interpretaciones contextualizadas y se debería debatir, incluso negociar, el informe final con el propio profesor.

La LOPEG, señala a los órganos colegiados de gobierno de los centros y a la inspección, como los agentes reconocidos para la evaluación del profesorado, hecho que contrasta con la opinión de los profesores, recogida en el estudio repetidas veces mencionado, en el que expresan su frontal rechazo a ser evaluados por los órganos colegiados del centro, por los padres y por los inspectores. Es la propia ley la que también reconoce el derecho de los profesores a participar en su propia evaluación. A nuestro juicio convendrá encontrar fórmulas de acción conjunta entre los órganos de poder y los profesores a fin de establecer estrategias evaluativas aceptadas por todos, la ley lo permite y el sentido común así lo aconseja.

CONSECUENCIAS EVALUATIVAS

Los propósitos y las consecuencias deben estar convenientemente relacionados, de esta forma se establece la coherencia global del modelo. Si proponíamos el desarrollo profesional como el propósito fundamental para la evaluación y sugeríamos la formación como el instrumento básico de intervención para alcanzarla, las consecuencias deberían orientarse a la consecución de dicho desarrollo profesional, que habitualmente se logra bien sea mediante el reconocimiento de los méritos profesionales contraídos o arbitrando los medios formativos que sean precisos para facilitar la mejora profesional.

La opinión del profesorado respecto a las consecuencias, (expresada en nuestro estudio, ya mencionado anteriormente) es totalmente clara. Consideran convenientes para el conjunto de la acción evaluadora, las consecuencias positivas, no necesariamente las que generan beneficios de tipo individual (p. ej. los sexenios), más bien prefieren las que pueda suponer ayudas económicas de tipo colectivo, o mejor aún las directamente ligada al desarrollo profesional como son: la promoción, o la consecución de tiempo o ayudas para la formación. Por el contrario muestran un rechazo claro a que la evaluación tenga efectos negativos.

LA FORMACIÓN CONTINUADA COMO MODELO DE INTERVENCIÓN

La evaluación moderna exige finalizar el proceso evaluador con un nuevo proceso orientado a la toma de decisiones. Toda evaluación de la que no se deriven consecuencias que faciliten la optimización del objeto/sujeto a evaluar pierde automáticamente todo sentido. No se trata de jugar a elaborar juicios de valor respecto de los demás, sino insertar el proceso evaluador en otro de orden superior que es el que mediante la adecuada toma de decisiones trata de servir desde la mejora de la calidad educativa, a la sociedad a la que se pertenece.

Ya hemos analizado en el apartado anterior las consecuencias que habitualmente emergen del proceso evaluador y volvemos a hacer referencia respecto de nuestra distinción entre la naturaleza del control y de la evaluación. Podemos observar, también, que la mayoría de las posibles consecuencias podríamos distribuirlas en dos grandes grupos: las que se podrían insertar en el marco de la acción controladora y las que podríamos situar en el de la cultura evaluativa.

La aplicación desculturalizada de premios y sanciones no aportan prácticamente nada en la mejora de la calidad de la enseñanza. La aplicación de los mismos principios, pero en el marco culturalizador del desarrollo profesional, aun siendo básicamente elementos claros de control, pueden conllevar mejora y crecimiento personal o colectivo, y aunque lo hagan de forma indirecta, mejorar la mencionada calidad de la enseñanza.

Sin embargo, la plataforma de intervención más sugerida (y sugerente) en la configuración de las praxis evaluativas actuales es la formación continuada. La evaluación en profundidad proporciona elementos racionales de diagnóstico, que permiten focalizar adecuadamente los procesos formativos individuales o grupales, para la mejora personal o institucional, condición indispensable (aunque no única) para la tan apreciada mejora de la calidad educativa. El desarrollo profesional otorga dirección y sentido al proceso a la par que coartada ética para su justificación. La acción armónica de los tres ingredientes, convierten a la evaluación de proceso técnico a instrumento culturalizador.

En la propia LOGSE se señala la importancia de la formación del profesorado como principio fundamental de intervención en la mejora de la calidad del sistema educativo, evidenciando por un lado la conjunción con la evaluación y el desarrollo profesional y por otro la estrecha conexión existente entre la formación del profesorado y las demandas sociales expresadas a través de sus respectivos sistemas educativos.

La formación continuada constituye en sí misma un modelo operativo completo y podríamos caracterizarla con Marcelo, C. (1989, p. 30) como «un proceso sistemático y organizado mediante el cual los profesores (en formación o en ejercicio) se implican individual o colectivamente en un proceso formativo que, de forma crítica y reflexiva, propicie la adquisición de conocimientos, destrezas y disposiciones que contribuyan al desarrollo de su competencia profesional».

Sparks y Loucks-Horsley (1990), (en Latorre 1994) plantean la existencia de cinco modelos o referentes de formación, ya que no los considera como propuestas cerradas

sino más bien como ámbitos de encuentro entre diferentes estrategias y actitudes y los enuncia y conceptualiza de la siguiente forma:

— *El modelo de desarrollo profesional orientado individualmente.* El proceso formativo que propone se caracteriza porque plantea la necesidad de que sean los propios profesores los que se responsabilicen de él y por tanto son ellos mismos los que determinan sus propios objetivos formativos y las actividades de formación que les permitirán alcanzarlos.

— *El modelo de observación-evaluación.* Este modelo parte del supuesto, que la reflexión y el análisis sobre la propia actuación docente son medios fundamentales para facilitar el desarrollo profesional.

— *El modelo de desarrollo y mejora de la enseñanza.* Este modelo parte del supuesto que es posible conseguir la mejora de la escuela y de la enseñanza si el colectivo de profesores se implica en tareas de desarrollo curricular y de la mejora de los procesos de enseñanza-aprendizaje. Implicación que supone que los profesores incorporan nuevas funciones orientadas a la toma de decisiones, a la mejora de la práctica docente a la innovación y al cambio curricular.

— *El modelo de entrenamiento.* En este modelo es el formador el que selecciona las actividades que han de ayudar a los docentes a lograr los resultados esperados. Se supone que los comportamientos o técnicas que implican estas actividades vienen avaladas por un buen número de investigaciones sobre la práctica de los docentes para una enseñanza eficaz. Los resultados esperados siempre tienen presentes la mejora de las concepciones y el pensamiento de los profesores.

— *El modelo indagativo.* Este modelo requiere una priorización por parte del profesorado de las necesidades, lo que se traduce en la identificación de un área de interés. Posteriormente se recoge información respecto de la misma y, basándose en la interpretación de estos datos, se realizan los cambios necesarios en la enseñanza. El nivel de implicación de este modelo en la práctica puede adoptar diversos niveles: individual, en grupos reducidos o de mayor tamaño, pudiendo llegar incluso a afectar a toda la institución.

ACCIONES EVALUATIVAS EN NUESTRO ENTORNO MÁS PRÓXIMO

Para poder analizar con cierta perspectiva las posibilidades de materializar modelos evaluativos con aplicación en la realidad, vamos a hacer un breve recorrido por las prácticas evaluativas más corrientes que se llevan a cabo en nuestro entorno más próximo.

Según la investigación realizada por González, M.C. y otros (1991, 1995) y tal como lo resumimos en Mateo, J. y otros (1995) y Escudero T. (1996), no existe en ningún país de la Unión Europea un solo plan global de evaluación de sus profesores (entendiendo este término en toda su amplitud). Sin embargo reconocen que si existen en estos países una evaluación administrativa periódica para los funcionarios, dentro de los que se incluyen los profesores.

Estas evaluaciones las suelen realizar los directores de los centros y/o los inspec-

res, su orientación es más de supervisión y control que de evaluación en sentido global, aunque en algún caso puede llegar a tener algún efecto en la promoción o retribución del profesorado.

Con los matices apuntados, Francia, Bélgica y Alemania, para profesores menores de 50 años, tienen una supervisión periódica y sistemática de su profesorado, llevada a cabo por directores escolares e inspectores.

Los objetivos perseguidos son variados, pero suelen estar relacionados con el control pedagógico, la planificación de efectivos, el cumplimiento, la aptitud y la dedicación profesional.

La evaluación es individualizada y se apoya en diversas fuentes como son los registros escolares, informes, entrevistas y visitas a clase. Esta evaluación no tiene consecuencias económicas para el profesor, aunque en Francia podría incidir en su ascenso o descenso en el escalafón. En estas evaluaciones el profesor recibe los informes y puede presentar alegaciones a los mismos.

Otros países de la Unión tienen también experiencias similares, aunque no tan generalizadas y sistemáticas. Así por ejemplo en Luxemburgo los inspectores controlan de esta manera el trabajo de los profesores de primaria, no el de los otros niveles. En Escocia existen ciertas experiencias piloto para establecer una evaluación sistemática a todo el profesorado, pero con una mayor orientación formativa.

La evaluación propiamente pedagógica, más centrada en la práctica docente y con orientación formativa, se va abriendo camino en muchos países, en coordinación o independientemente de la evaluación más controladora del cumplimiento docente. En Inglaterra y Alemania existen bastantes experiencias en este sentido, con vocación de extenderse de manera generalizada.

Lo que sí existen en todos los países son procesos de evaluación de los profesores más específicos, en momentos concretos o para determinados fines selectivos y promocionales.

En el Estado Español al igual que en los países de nuestro entorno, tampoco se puede decir que exista un plan sistemático y global de evaluación del profesorado de primaria y secundaria en el sentido aludido anteriormente.

Obviamos aquí el reconocimiento que desde la década de los ochenta se hace de las actividades de formación del profesorado no universitario y de las actividades docentes e investigadoras del profesorado universitario y que se ven incentivadas económicamente con unos complementos de sueldo denominados sexenios en el primer caso y tramos en el segundo.

Sin embargo también existe aquí un cierto movimiento hacia la evaluación del profesorado más pedagógica en términos similares a los observados en la Unión Europea.

La reciente Ley de la participación, la evaluación y el gobierno de los centros (LOPEG), que tiene rango de orgánica, contempla, tal como ya hemos visto, en su título tercero, la evaluación de la función docente, estando, tal como ya se comentó, encomendada a la Inspección Educativa y a la Dirección de los Centros la responsabilidad básica de llevarla a término. Sin embargo la Ley aún está pendiente de ser desarrollada y normativizada por las diferentes administraciones autonómicas.

PRINCIPIOS Y CONDICIONES PARA LA CONFIGURACIÓN DE UN MODELO DE EVALUACIÓN DEL PROFESORADO

Hemos comprobado que no existen, al menos en el ámbito europeo, experiencias completas y lo suficientemente prolongadas, como para poder extraer conclusiones definitivas acerca de las características posibilidades y límites, que deberían configurar un modelo operativo de evaluación del profesorado, pero si creemos que se han llevado a cabo suficientes aproximaciones como para intuir el camino a seguir, para dotar a la evaluación de un punto de arranque con perspectivas de implantación. Elementos que aportamos aquí a modo de cierre del itinerario que proponíamos al principio.

En primer lugar consideramos que el marco que ofrece la ley orgánica 9/90 puede abrir expectativas fundadamente optimistas, si se aprovechan las experiencias anteriormente mencionadas y se respetan algunos principios sobre los que existe consenso en cuanto a su capacidad de crear unas condiciones iniciales favorables a la generación de cultura evaluativa.

Del análisis de las diferentes propuestas sugeriríamos respetar Al menos las siguientes recomendaciones técnicas:

a) Partir siempre de la existencia de un marco legal que regule y normativice el hecho evaluativo. Su existencia supone, entre otras cosas, el constatar el derecho que asiste a la sociedad a tener un profesorado competente al servicio del sistema educativo y al profesor de disponer de un marco de garantías que protejan sus derechos.

b) Profundizar en la dirección de modelos evaluativos participativos, donde el profesor, tal como ya preve la ley, sea sujeto activo de su propia evaluación.

c) Conectar la evaluación del profesorado a los procesos de formación, de mejora y desarrollo profesional y orientada a los procesos de innovación y cambio institucional.

d) La política evaluadora debe ser clara, conocida y debe tratar de gozar de una amplia aceptación, por lo que requiere mínimamente que el modelo:

- fije en cada caso los contenidos a evaluar.
- determine el propósito de la evaluación.
- señale los agentes evaluadores.
- seleccione las estrategias más adecuadas.
- utilice técnicas rigurosas de análisis e interpretación de la información.
- establezca el tipo de ayudas y consecuencias que se derivarán de la evaluación.
- pacte los modelos de formación y de retroalimentación.

Como recomendación general hay que dejar tiempo al profesorado para que vaya incorporando los elementos de decisión y de mejora surgidos de su evaluación, sin agobiarle con continuos procesos evaluativos. La marcha evaluativa debe hacerse por expansiones concéntricas que van avanzando lenta pero inexorablemente.

Es muy importante conseguir y mantener en todo momento la implicación del profesorado e incrementar su convicción en cuanto a sus posibilidades de crecimiento

y desarrollo tanto personal como institucional y su confianza en los procesos de formación como modelo general de innovación y de cambio.

La administración por su lado debe ofertarle los recursos necesarios para hacer posible el proceso de mejora. En este esquema lógicamente la sanción debe contemplarse como una posibilidad que si bien existe, se sitúa bajo el signo de la excepcionalidad.

Finalizamos este apartado sintetizando las conclusiones a las que llegamos a partir del trabajo de investigación acerca de la opinión de los profesores sobre su propia evaluación (Mateo y otros. 1995). Opiniones que deberán ser tenidas en cuenta, no tanto como elementos determinantes, sino más bien, como condiciones a debatir y considerar a fin de favorecer la generación de un clima positivo, para conseguir que el proceso evaluador supere la cota de lo cualidad puramente técnica y sea también un proceso culturalizador.

a) Los profesores consideran necesaria su evaluación pero entienden que ésta debe formar parte de la evaluación global del sistema.

b) Los profesores opinan claramente que la evaluación debe tener objetivos formativos para el individuo y para el sistema, mientras que mantienen reservas respecto a conectar la evaluación a objetivos de control o de incentivación económica personal.

c) La clase y las actividades directamente relacionadas con ella son los factores que los profesores consideran más importantes para su evaluación. El resto de actividades, incluidas las de formación y servicio, son poco consideradas por los profesores como elementos a tener en cuenta para su evaluación.

d) Rechazan el ser evaluados por agentes externos: sindicatos, padres e inspección. Hay mayor aceptación respecto de agentes próximos al profesor (alumnos, compañeros...).

e) No aceptan los efectos de la evaluación de tipo negativo y valoran escasamente los efectos económicos de tipo individual. Valoran mejor los efectos de promoción y ayudas a proyectos de formación e innovación.

f) Piden clarificación en los objetivos de la evaluación

g) Exigen la confidencialidad de la información evaluativa.

A MODO DE SÍNTESIS

La evaluación del profesorado es un factor esencial en cualquier propuesta de mejora del sistema educativo, sin embargo su operativización es altamente problemática y aún quedan por llegar soluciones satisfactorias en la mayoría de los países de nuestro entorno.

En nuestro país la LOGSE en su artículo 62 ya preveía su evaluación, posteriormente la reciente LOPEG caracteriza en su título III con más detalle el sentido de la misma, señalando específicamente al servicio de Inspección como encargado principal de llevarla a cabo, garantizando en todo caso la participación del profesorado en su propia evaluación.

Dos propósitos fundamentales justifican prioritariamente la evaluación del profe-

sorado, la mejora de la calidad de la enseñanza y el desarrollo profesional del docente. Se entiende que ambos propósitos van íntimamente unidos, así consideramos que el desarrollo profesional entendido armonizado en sus dimensiones individual e institucional concluye necesariamente en la mejora de la calidad de la docencia.

Los modelos evaluativos actuales se construyen a partir de configurar y engarzar elementos de tipo estructural y metodológico. Para diseñar un modelo, desde la perspectiva estructural hay que decidir el propósito (desarrollo profesional...), la orientación (formativa, sumativa...) los conocimientos y competencias a evaluar (dominio de la materia, competencias instruccionales, competencias de evaluación...) los elementos de profesionalidad (servicios a la comunidad educativa...).

Desde la metodológica habrá que decidir en torno a otra serie de ingredientes, como son las fuentes de información, los agentes evaluativos, los procedimientos y estrategias de recogida de información y los procesos de análisis de la misma, así como la orientación paradigmática general que afecta a todos los componentes, pero muy especialmente al enfoque en los procesos de toma de decisiones.

El modelo es la resultante en la práctica de la construcción anterior. En la actualidad, las administraciones educativas de los países más avanzados, están claramente evolucionando desde los modelos evaluativos concebidos fundamentalmente como instrumentos de control, en ocasiones ligados a mejoras salariales (pago por méritos), a modelos que entienden la evaluación como un proceso de desarrollo profesional (individual e institucional) asociado a la formación y perfeccionamiento del profesorado.

Se constituye así el eje formado por la evaluación, el desarrollo profesional y la formación continuada como vertebrador de la mayoría de las modernas concepciones evaluativas de los docentes.

Se perfila por tanto, tal como reconoce la propia LOGSE, la importancia de la formación continuada del profesorado, como elemento fundamental en la mejora de la calidad del sistema educativo y por tanto se reafirma su significación en el eje anteriormente mencionado.

Sparks y Loucks-Horsley nos plantean cinco modelos o referentes de formación permanente que cubren prácticamente todo el espectro de posibilidades formativas. A nuestro juicio se deben primar aquellos que favorezcan simultáneamente al sujeto y a la institución, que comporten mayor nivel de implicación y compromiso del docente con el centro y que conceptualicen al profesor como profesional reflexivo y crítico capaz de adoptar un papel activo en su propio proceso de aprendizaje.

La importancia de la formación continuada, se traduce en la creciente atención que todas las Administraciones Educativas le conceden. Existe una amplia oferta formativa a nivel de todo el estado y de las autonomías con competencias educativas, que va desde la formación dirigida a la generalidad de los docentes y pensada básicamente como instrumento de adaptación a los cambios del sistema educativo, hasta planteamientos más institucionales dirigidos a cubrir necesidades a nivel de un centro tomado en su conjunto.

Con la evaluación del profesorado no se agotan las posibilidades evaluativas del personal docente; en los sistemas educativos más desarrollados se admite cada vez

más, la relevancia de la función directiva, por su clara incidencia en la calidad del sistema escolar en particular y del educativo en general. De aquí el esfuerzo dedicado a los procesos de selección, formación y evaluación de los directivos de centros escolares.

En nuestro sistema normativo y tal como se señala en la LOGSE, los Directores de los centros sostenidos por fondos públicos, son elegidos democráticamente por los Consejos Escolares. Sin embargo es preceptivo para ser candidato a esta elección, el estar previamente acreditados. Las condiciones de acreditación están recogidas en el artículo 19 de la LOPEG. Las funciones y competencias que le corresponden al Director se definen en el artículo 21 de la misma ley y finalmente en el 34 la obligatoriedad de las administraciones educativas de establecer planes de evaluación de la función directiva.

Finalizamos señalando, que cualquier modelo evaluativo que se pretenda implantar, deberá partir de una conceptualización contextualizada en la realidad sobre la que pretende incidir. Deberá conjugar rigor y funcionalidad. Atrevimiento y prudencia. Esmero en el trato respecto de los elementos personales e institucionales a los que afecta e intervienen en el modelo. Incidir en lo formativo. Cuidadoso en lo legal y lo ético, deberá en definitiva superar la frontera de lo técnico para sumergirse en el ámbito de lo cultural.

REFERENCIAS BIBLIOGRÁFICAS

- DENHAM, C.H. (1987). A perspective on the major purposes and basic procedures for teacher evaluation. *Journal of Personnel Evaluation in Education*. 1, 1, 29-32.
- ESCUADERO, T. (1996). Proyecto docente de Cátedra de Universidad no publicado. Zaragoza: Universidad de Zaragoza.
- GAIRIN, J. y DARDER, P. (1994). *Organización de Centros Educativos. Aspectos básicos*. Barcelona: Praxis.
- GIMENO, J. (1993). El profesorado ¿Mejora de la calidad o incremento del control? *Cuadernos de Pedagogía*, 219, 22-28.
- GONZÁLEZ, M.C. y otros. (1992). *La evaluación de profesores*. Proyecto de investigación. Madrid: CIDE.
- GONZÁLEZ, M.C. y otros (1995). *Evaluación del profesorado de educación secundaria. Análisis de tendencias y diseño de un plan de evaluación*. Madrid: MEC-CIDE.
- HARRIS, B.M. (1986). *Development teacher evaluation*. Boston. Ma.: Allyn and Bacon.
- LATORRE, A. (1994). Models de desenvolupament professional del professor. *Temps de Educació*, 11, 59-87, 1^{er}. semestre 94.
- MATEO, J. (1993). *La evaluación del profesorado de secundaria*. Ponencia en el seminario sobre las reformas educativas actuales en España. Su evaluación: metodología y resultados, 26-28 octubre. Madrid: CIDE-UNED-CDLM.
- MATEO, J. (1995). *Avaluació*. Document intern. Barcelona: Generalitat de Catalunya. Departament d'Ensenyament.
- MATEO, J. y otros (1995). *La evaluación del profesorado de secundaria: la opinión del profesor*. Proyecto de investigación del CIDE. Memoria Final. Barcelona.

- RIEGLE, R.P. (1987). Conceptions of faculty development. *Educational Theory*, 37, 53-59.
- RUL, J. (1993). *La organización escolar y la inspección educativa. Una inspección educativa como factor de evaluación sistemática*. Conferencia. Actas II Congreso de Inspectores de Educación. Alicante.
- SCHALOCK, M.D. y Otros, (1993). Teacher productivity revised: Definition, theory, measurement and application. *Journal of Personnel Evaluation in Education*, 8, 2, 179-196.
- SCRIVEN, M. (1988). Duty-based teacher evaluation. *Journal of Personnel Evaluation in Education*, 1, 4, 319-334.
- SCRIVEN, M. (1994). Duties of the teacher. *Journal of Personnel Evaluation in Education*, 8, 2, 151-184.
- SPARKS, D. y LOUCKS-HORLEY, S. (1990). *Five models of staff development*. Oxford, OH: National Staff Development.

LAS AYUDAS DEL PROFESOR EN EL APRENDIZAJE COOPERATIVO Y SU INFLUENCIA EN LA EVALUACIÓN CRITERIAL

Leonor Buendía Eisman
Universidad de Granada

RESUMEN

Con esta investigación pretendemos conocer las ayudas que los profesores proporcionan a los alumnos cuando trabajan cooperativamente, y como influyen ambas variables en el nivel de logro de los objetivos que pretendemos conseguir en primer curso de Educación Primaria. La evaluación de los alumnos se ha realizado con pruebas criterioles, elaboradas por los propios profesores que han participado en la experiencia; y la actuación docente o ayudas que el profesor proporciona, con una escala de observación, realizada para este fin.

ABSTRACT

With this investigation we intend to know the helps that teachers provide to the pupils in cooperative work and the influence of both variable in the objectives that we intend to obtain in first Primary Education course. The evaluation of the pupils has been accomplished with tests criterion-referenced, elaborated by the own teachers that have participated in the experience, and the help that the teacher provides, through an observation scale accomplished for this purpose.

1. INTRODUCCIÓN

En los últimos años uno de nuestros focos de atención ha sido la evaluación de los aprendizajes en el aula y los factores que colaboran o no a ello, (Buendía, 1985, 1987, 1993; Buendía y Fernández, 1991), en la línea de (Bennet, Deforges, Cockburn y Wilkinson, 1984; Bennet y Kell, 1988; Dunne y Bennett, 1990). A través de la investigación hemos podido constatar el importante papel de las interacciones en el aula sobre las actividades cognitivas y sociales. Hablar de educación, supone en cierto sentido, hablar de comunicación (Myers y Lamm, 1976; Wittrock, 1974); sin embargo estamos convencidos de que el análisis de la interacción no es el único análisis pertinente para una mejor comprensión del proceso de enseñanza-aprendizaje; más bien, el análisis debe ampliarse al estudio de la **interactividad** profesor/alumno, entendiendo por esta aquella situación en la cual es posible articular las actividades del profesor y de los alumnos en torno a un contenido o a una tarea de aprendizaje. Esto implica prestar una atención especial a la forma de trabajar los alumnos en estrecha vinculación con las actuaciones del profesor. La articulación conjunta de los tres elementos claves en la situación educativa: profesor, contenidos y alumnos, ha sido objeto de interesantes estudios cualitativos en los últimos años (Buendía, 1993, Pegalajar, 1995). En este sentido Coll y colaboradores, (1992:192) apuntan que la interactividad profesor/alumnos, entendida como las formas de organización de la actividad conjunta, define el marco en el que cobran sentido, desde el punto de vista de la influencia educativa, las actuaciones respectivas y articuladas de los participantes en el transcurso de un proceso de enseñanza-aprendizaje. En este trabajo hemos retomado el tema, integrando otras variables, e intentando implicar a los profesores tanto en el conocimiento de su práctica como profesores como en la confección de instrumentos de evaluación, acordes con los criterios de evaluación que ellos mismos establecen.

2. OBJETIVOS E HIPÓTESIS DE INVESTIGACIÓN

Nos planteamos dos objetivos generales. El primero tiene un **carácter exploratorio** y se concreta en la observación, registro, descripción y análisis de las ayudas proporcionadas por el profesor.

En segundo lugar pretendemos conocer como influyen estas ayudas y la metodología de trabajo que sigue el alumno, en el nivel de logro de cada uno de los objetivos didácticos que para un proyecto educativo se propone el profesor.

Nuestra hipótesis de partida es que los alumnos que adquieren los conocimientos en una actuación compartida con sus compañeros (aprendizaje cooperativo) y con una ayuda pedagógica (actuación docente) 'significativa', consiguen mejores niveles de logro, en los objetivos didácticos de un proyecto educativo, que aquellos cuyo trabajo se enmarca en una actuación individual sin que realmente exista ayuda significativa.

3. VARIABLES

Las variables independientes son:

1. Características personales de los alumnos.
2. Actuación docente del profesor.
3. Metodología de trabajo del alumno.
4. Variable dependiente: Nivel de logro de los objetivos didácticos en Matemáticas.

3.1. Características personales de los alumnos

En la evaluación inicial comenzamos conociendo las aptitudes de los alumnos, referidas a madurez intelectual, atención, razonamiento, etc.

Para medirlas disponemos de diferentes pruebas estandarizadas, habiéndonos decidido por la batería de aptitudes diferenciales y generales (BADIG) de C. Yuste (1984) para los alumnos de Primaria. Aunque la prueba en su totalidad está concebida en seis niveles para atender las diferentes etapas, hemos seccionado sólo las partes A, B, y C que están dirigidas a niños de 4 a 9 años. La características de la prueba, por las que hemos realizado la opción son:

- a) Son pruebas baremadas en España.
- b) El factor tiempo no tiene importancia en ninguna prueba.
- c) Las pruebas elegidas para Primaria han sido específicamente preparadas para alumnos de 6 a 9 años. (Opción B y C).
- d) Las diferentes partes se pueden aplicar por separado permitiéndonos la selección de aquellas que más interesan para nuestro estudio, al presentar todas, individualmente tomadas, buenos índices de fiabilidad y validez.
- e) Utiliza un material fácilmente representable y manejable para el niño.

En todas las pruebas se usa casi exclusivamente el dibujo como elemento de representación conceptual no insinuando una lógica formal más que en seis ítems del test de habilidad mental-no verbal, y aún así muy elemental.

La prueba mide los siguientes factores:

- a) Madurez intelectual verbal (I.G.V.).
- b) Madurez intelectual no verbal (I.G.N.V.).
- c) Ordenes Verbales (O.V.).
- d) Aptitud para el cálculo.
- e) Habilidad Mental No Verbal. (H.M.N.V.).
- f) Razonamiento con Figuras (R.L.).
- g) Atención-Observación (At.).

Hemos considerado pertinente confeccionar un baremo propio a partir de la muestra que hemos seleccionado para la presente investigación. Esta muestra consta de 12 grupos-clase de 1º de E.G.B. de la provincia de Granada. Este baremo nos permitirá situar a cada sujeto en cualquiera de los factores medidos en relación a los demás sujetos, posición que vendrá dada por el percentil que corresponda a cada puntuación directa.

3.2. Actuación docente

Del análisis de contenido de los diseños curriculares se obtuvieron las cuatro categorías de actuación docente que, según el marco teórico de partida, permitirían una ayuda significativa. Éstas son:

- 1º) El profesor deberá planificar y organizar las distintas relaciones que se pretenden establecer en su grupo de alumnos.
- 2º) El maestro constituye un modelo adulto significativo para su grupo de alumnos.
- 3º) El maestro tiene una importante función en cuanto a la formación en la clase de un grupo humano cohesionado.
- 4º) El maestro tiene un papel esencial como mediador de aprendizajes significativos.

Definición de las categorías:

a) Cada profesor deberá planificar y organizar las distintas relaciones que se pretenden establecer en su grupo de alumnos.

Dicha planificación requiere un estudio en profundidad de las características particulares que concurren en esa comunidad educativa, desarrollando estrategias que posibiliten unas relaciones entre **todos** sus elementos, lo más ricas posibles. Hace referencia a: Relaciones entre los alumnos, relaciones de los alumnos con los profesores, relaciones con el resto de la comunidad en la que se ubica la escuela, etc...

b) El maestro constituye un modelo adulto significativo para su grupo de alumnos.

No podemos considerar al maestro como un individuo más del grupo y en igualdad con los restantes miembros. Su intervención en el mismo, sus actitudes, su emotividad... conforman para los alumnos un modelo significativo que, junto con el de sus padres y otros adultos relevantes en su vida, contribuyen a forjar una imagen adulta que en buena medida va a incidir en su desarrollo.

El maestro, desde esta perspectiva deberá desarrollar una serie de actitudes tales como: Actitud afectuosa, actitud no autoritaria, actitud de respeto y confianza en el niño, mantener conductas coherentes, etc.

c) El maestro tiene una importante función en cuanto a la formación en la clase de un grupo humano cohesionado.

Uno de los objetivos de la educación, es potenciar el desarrollo del niño en su dimensión social; en este sentido la escuela ofrece un marco inigualable en donde este desarrollo puede favorecerse. En la clase debe establecerse una red de comunicación

multidireccional entre profesor-grupo de alumnos, el grupo de niños entre sí, el profesor-alumno, profesor-pequeño grupo, alumno-alumno, etc...

Esta red de comunicación la hace posible el profesor a través de distintos recursos, entre los que podemos destacar: favorecer la comunicación informal, creación de espacios que estimulen diversas formas de agrupamiento en contextos variados, hacer de la clase una asamblea de discusión y opinión, etc.

d) El maestro tiene un papel esencial como mediador de aprendizajes significativos para los alumnos.

En la medida que el profesor conoce los procesos de adquisición de conocimiento y las características evolutivas de sus alumnos, diseña su intervención didáctica y promueve la construcción de aprendizajes significativos. Interviene estableciendo conexiones entre los conocimientos nuevos y los que el niño ya tiene, dedicando especial atención a los aspectos que parecen no estar claros, que son incongruentes o sobre los que no existe acuerdo en clase, planteando problemas, diálogos, etc...

3.2.1. Construcción de la lista de rasgos de observación sobre actuación docente (ayudas que proporciona el profesor)

En la literatura existente, hay un elevado número de estudios sobre interacción profesor-alumno, con elaborados sistemas de categorías para la observación de la actuación del profesor o de las interacciones en el aula (Good y Brophy; Flanders, Landsheere,...).

Sin embargo, la actuación del profesor, bajo la concepción constructiva de la enseñanza-aprendizaje, lejos de ser planteada como un conjunto de reglas y procedimientos que se aplican mecánicamente, lo cual daría lugar a una actuación determinada previamente, es presentada como una «herramienta de reflexión» para que los profesores, según los problemas que encuentren al enseñar a los alumnos, presenten un modelo u otro de actuación.

Por esta razón nos ha parecido tan importante elaborar nuestra propia lista de rasgos sobre la actuación docente, extraídos de la observación directa de los profesores en sus aulas.

El proceso de observación exploratoria, tuvo lugar en el mes de octubre, durante una semana, diariamente 45 minutos. Tres clases elegidas al azar de las que han participado en la experiencia, fueron grabadas en vídeo por un observador. Cada día se grabó la actuación del profesor en un área distinta, siguiendo la siguiente secuencia:

Lunes	Matemáticas	(45')
Martes	Lengua	(45')
Miércoles	Sociales	(45')
Jueves	Matemáticas	(45')
Viernes	Lengua	(45')

Así pues se analizaron 5 cintas de vídeo de 45 minutos cada una por clase observada. El total de cintas analizadas fueron:

	<i>Tiempo</i>
6 cintas de Matemáticas	4 h. 30'
6 cintas de Lengua	4 h. 30'
3 cintas de Sociales	2 h. 15'

Del análisis de estas cintas obtuvimos un catálogo de estrategias, actitudes y ayudas que utilizan los profesores de Primero de Educación Primaria.

Sólo seleccionamos los rasgos que respondían a tres categorías de las cuatro extraídas de análisis del diseños curricular, expuestas anteriormente, cuando se definió la variable de «actuación docente».

De la primera categoría hemos prescindido, por no considerarla relevante para nuestro trabajo, al tratarse del papel del profesor y alumnos con otros miembros de la comunidad educativa.

Una vez extraído este primer listado se procedió a agrupar cada rasgo en la categoría correspondiente.

Por ejemplo según los análisis de los documentos, el profesor debe ser mediador de aprendizajes significativos (Diseño curricular de primaria, p. 104). Textualmente dice:

«El profesor ocupa un lugar importante en la interacción educativa. En la medida que conoce los procesos de adquisición de conocimiento y las características evolutivas de sus alumnos, diseña su intervención didáctica para promover la construcción de aprendizajes significativos..., plantea problemas, orienta la conversación, etc...».

Con este texto de referencia todas las observaciones hechas sobre el profesor como mediador de aprendizajes significativos entran a formar parte de esta categoría.

Así se continuó adscribiendo cada uno de los rasgos a las tres categorías comentadas.

Se observaron estrategias, comportamientos y actitudes en los profesores cuya presencia era negativa para una ayuda significativa. Estas conductas también han sido incluidas pero con una formulación positiva. Por ejemplo, si hemos observado una actitud racista, como rasgo en la categoría (modelo adulto significativo) se ha planteado como «no racista ni sexista», o si se observa que el profesor impide la participación del alumno cuando está explicando un concepto, lo hemos plasmado como: «*anima a intervenir*». La adscripción de los rasgos a las diferentes categorías fueron realizadas por dos observadores, los mismos que previamente habían visualizado las cintas y obtenido las listas de rasgos.

Terminada la adscripción de rasgos y tras sucesivas revisiones el listado quedó finalmente como sigue:

A. Sirve de modelo adulto significativo.

Entendemos que el maestro sirve de modelo significativo adulto cuando fomenta una serie de conductas tales como:

1. Respeto y confianza en el niño.
2. Seguridad en sí mismo.
3. Conducta no autoritaria.
4. Manifiesta opiniones que sirven para combatir estereotipos sociales.

B. Crea un grupo humano cohesionado.

El profesor crea un grupo humano cohesionado cuando:

5. Establece una red de comunicación multidireccional.
6. Favorece la comunicación informal.
7. Crea un clima afectivo positivo.
8. Planifica e interviene en el trabajo grupal de clase.
9. Establece relación con distintas personas que entren en el aula.
10. Establece relaciones entre los grupos.

C. Mediador de aprendizajes significativos.

El profesor es mediador de aprendizajes significativos cuando:

11. Plantea problemas.
12. Establece conexiones entre el conocimiento nuevo y el que el niño ya tenía.
13. Ofrece información relevante.
14. Observa el proceso y se detiene en aquellos aspectos que parecen no estar claros.
15. Crea situaciones que permiten al alumno reflexionar sobre su propio proceso de aprendizaje.
16. Organiza debates.
17. Orienta al diálogo.
18. Ayuda a los que lo necesitan a expresar sus ideas.
19. Anima a intervenir.

Cada uno de estos rasgos fueron suficientemente definidos y operativizados para la formación de los observadores. Las definiciones de cada rasgo fueron entregadas, junto con la hoja de registro (ver anexo) y las aclaraciones pertinentes, a los observadores de las aulas.

Procedimiento seguido por el observador

El observador anotará:

- (+) Presencia de la conducta.
- (-) Presencia de la conducta contraria a la contemplada en la escala.
- (A) Cuando la actuación docente va dirigida a la clase en general.

(G) Cuando la actuación docente va dirigida a algún grupo de trabajo cooperativo o colaborativo.

(I) Cuando la actuación docente va dirigida a un alumno en particular.

Ejemplo: En el rasgo C.19 nos encontramos el siguiente registro:

+	G
---	---

Esto significa que el maestro ha estado animando a intervenir a los alumnos de algún grupo de aprendizaje. (Ver anexo con la lista de rasgos).

Cada cinco minutos el observador anota las conductas que aparecen en un cuadro distinto. Cada columna de celdillas equivale a un momento de observación en el periodo de 5 minutos. Al final de la sesión de observación podremos contabilizar la frecuencia de dicha conducta y si se ha realizado con la clase en conjunto, con los grupos pequeños de trabajo o con sujetos individualmente.

Características técnicas de la escala

La adscripción de los rasgos a las diferentes categorías fue realizada, como decimos anteriormente, por dos observadores, los mismos que visualizaron las cintas grabadas de las clases.

La fiabilidad de la escala de observación se obtuvo mediante el porcentaje de acuerdo.

$$C\% A = \frac{Na}{Na + Nd} \times 100$$

Siendo: C% A: Porcentaje de acuerdo; Na: Número de acuerdos; Nd: Número de desacuerdos.

El porcentaje obtenido fue del 94% considerando, en consecuencia, que el resultado era bueno y que existía concordancia en la adscripción de los rasgos a la conceptualización (categoría) que habíamos realizado sobre los tres conjuntos de rasgos.

Somos conscientes de las limitaciones de este índice (Hartman, 1979; Bijou, Peterson, Harris, Allen and Johnston, 1969; Hopkins y Hermann, 1977) para el cálculo de la fiabilidad de las observaciones, pero en este caso sólo pretendíamos conocer el grado de acuerdo en la adscripción de los rasgos a las categorías, por eso prescindimos de otros índices más elaborados como el índice de concordancia ponderada (Harris y Lahey, 1978), o el índice de concordancia en su forma canónica y matricial (Krippendorff, 1980) para tres observadores.

Además nos interesaba conocer la fiabilidad de las puntuaciones como índice de

constancia o estabilidad. Para ello correlacionamos las puntuaciones de cada sesión de observación para cada clase observada. Los índices de correlación estuvieron todos por encima de 0.70.

3.3. Metodología de trabajo de los alumnos

En las situaciones escolares, las relaciones con los compañeros pueden ser estructuradas para crear una interdependencia importante a través del aprendizaje cooperativo.

La característica clave que distingue las situaciones cooperativas de otras situaciones de aprendizaje, es la interacción entre estudiantes.

La influencia de los procesos interactivos, tanto entre iguales como entre el profesor y alumnos, se ha estudiado desde diferentes perspectivas: Explicación del progreso del sujeto a través del conflicto socio-cognitivo (Doisé, Mugny, y Perret-Clermont, 1975); reestructuración intelectual a través de la regulación recíproca que ejercen entre sí los alumnos cuando interactúan, y la posterior interiorización que lleva a la autorregulación individual (Forman y Cazden, 1984); la interacción social como origen y motor del aprendizaje, (Vygotsky, 1973); el aprendizaje como actividad comunal y un modo de compartir la cultura (Bruner, 1986).

Desde una u otra perspectiva se aboga por un trabajo en las aulas diferente a la actuación en solitario que con frecuencia la escuela ha fomentado. Cuando los profesores aluden al trabajo en equipo, a menudo se refieren al grupo de niños que se sientan juntos pero que trabajan en tareas individuales.

La falta de oportunidades para que los alumnos interactúen en el aula está bien documentada (Galton, Simon y Croll, 1980; Gerleman, 1987; Goodlad, 1984; Lockeed y Harris, 1984, Graybeal y Stodolsky, 1985). Cuando trabajan en grupos raramente funcionan como tales y por lo tanto los niveles de cooperación son bajos. Es necesario un cambio metodológico que permita mayor interacción entre los alumnos.

Una de las mayores justificaciones del trabajo cooperativo se encuentra en el estudio de Allen y Feldman (1973). Estos autores comprobaron que las dificultades de determinados conceptos solo son percibidos por los alumnos, no así por el profesor. En este estudio se visualizaron vídeos con dos tipos de lecciones (fáciles y difíciles), mientras que los alumnos si se percataron de esta diferencia, no ocurrió igual para los profesores que sobreestimaron la comprensión de los alumnos en la lección difícil. Según esto, los alumnos que tienen dificultades mientras aprenden, podrán beneficiarse mejor trabajando con otros estudiantes.

Aunque todos parecen estar de acuerdo en el importante papel jugado por la interacción social, no hay la misma consistencia en los resultados cuando se comparan los logros entre diferentes métodos de aprendizaje en régimen cooperativo (Webb, 1984: 159). La literatura sugiere que es importante observar y describir las diferentes formas de aprender en pequeño grupo y entender sus puntos fuertes y limitaciones en relación a diferentes metas de enseñanza. (Good et al., 1990).

Las mayores inconsistencias en los resultados se han explicado en base a las diferentes técnicas de aprendizaje cooperativo (Slavin, 1989) y la materia a estudiar (Sharan, 1980).

En función de esto, nosotros hemos diferenciado entre:

1. Aprendizaje individualizado
2. Aprendizaje cooperativo
3. Aprendizaje colaborativo

En una estructura **individual** los alumnos son recompensados sobre la base de la calidad de su propio trabajo, con independencia del trabajo de los otros participantes.

Se da estructura **cooperativa** cuando los objetivos que persiguen los participantes estén estrechamente vinculados entre sí, de tal manera que cada uno de ellos puede alcanzar sus objetivos sí, y sólo sí, los otros alcanzan los suyos. La técnica de trabajo seguida fue «*Jigsaw*» o **técnica de rompecabezas**. (Aronson y Bridgeman, 1981; Aronson y otros, 1978; Aronson y Osherow, 1980).

En una estructura **colaborativa** los participantes trabajan juntos durante todo el tiempo de la ejecución de la tarea, en lugar de hacerlo individualmente en cada una de los diferentes componentes de la misma.

La técnica de trabajo seguida fue «**Learning Together**» (Johnson y Johnson, 1975).

3.4. Nivel de Logro

Entendemos por nivel de logro el grado en el que los sujetos llegan a conseguir los objetivos didácticos, en las diferentes evaluaciones, que para cada proyecto de trabajo están planteados.

En este caso nos referimos al grado en el que los alumnos consiguen superar el nivel propuesto para el tercer proyecto de Matemáticas.

Los contenidos de este proyecto de trabajo para matemáticas fueron:

1. Conocimiento de los números naturales hasta el 7.
 - 1.1. Repaso de los números del 1 al 5 (Proyecto nº 2).
 - 1.2. Presentación del número 6.
 - 1.3. Presentación del número 7.
2. Sumas y restas en campo numérico hasta el 7.
3. Concepto espacial cerca/lejos, dentro/fuera y alrededor.
4. Interpretación de datos estadísticos: más que, menos que y más repetido.

Este material se la ha facilitado a los 12 profesores para que fueran los mismos contenidos los que se trabajaran en el aula, después de una reunión conjunta y coincidir en el momento aproximado de programación en el que se encontraban. No comenzaron el proyecto todos el mismo día, sino que dependió del ritmo que cada maestro llevaba en el desarrollo normal de su programación.

3.4.1. *Elaboración de la prueba criterial para la evaluación*

La finalidad de la **formación matemática**, es la de favorecer, fomentar y desarrollar

en los alumnos la capacidad para explorar, formular hipótesis y razonar lógicamente, así como la facultad de usar de forma efectiva diversas estrategias y procedimientos matemáticos para plantearse y resolver problemas relacionados con la vida cultural, social y laboral.

La concepción del área matemática como integradora y cultural, y la priorización de las formas empíricas e inductivas de adquisición del conocimiento, nos han servido como referente para desarrollarla.

Elaboramos una prueba criterial para evaluar el grado de superación de los objetivos de matemáticas y además, el nivel individual de desarrollo de cada uno de los alumnos respecto al punto inicial de partida.

El proceso consta de las fases siguientes:

1. Validación de los contenidos propuestos por el Ministerio con los textos de las distintas editoriales que tienen propuestas en estos niveles.
2. Relación de los objetivos de matemáticas de este nivel con los objetivos generales de área, con los objetivos de primaria y con los objetivos de ciclo.
3. Especificación del dominio de aprendizaje.
4. Elaboración de la prueba, tomando como referencia los ítems redactados por los textos de las diferentes editoriales.
5. Valoración de la prueba por un grupo de expertos.
6. Corrección y posterior elaboración de la prueba definitiva.
7. Aplicación.

3.4.2. *Establecimiento del dominio de aprendizaje (validación de contenidos y propuesta de objetivos)*

Determinar el dominio de aprendizaje, objeto de evaluación, supone definir claramente el propósito de la prueba, consensuar un listado de objetivos que se demuestren como relevantes y especificar el dominio, de forma que este oriente claramente la redacción de ítems.

Llegamos a su determinación mediante una estrategia lógico-racional realizando un análisis exhaustivo de las propuestas de seis editoriales (Alhambra-logman, Anaya, Santillana, S.M., Bruño, Luis Vives) y el refrendo que a nivel legal y científico tienen estas propuestas.

Determinada la relevancia de las metas en cualquiera de los libros de recursos utilizados por los profesores, nos quedaba hacer una selección de objetivos para el citado proyecto. Llegamos a la siguiente conclusión mediante el consenso entre los profesores.

Objetivos generales de Área para la primera quincena de diciembre

A) Números y operaciones.

M-1 Conocer los principales sistemas de numeración.

M-2 Conocer y manejar operaciones elementales de cálculo: suma y resta.

- M-3 Utilizar la operativa para resolver situaciones problema de la vida cotidiana.
- M-4 Efectuar operaciones sencillas de cálculo mental, mostrando confianza en las propias capacidades.
- B) Formas geométricas y situación en el espacio.
- M-6 Lograr un progresivo dominio de los conceptos espaciales.
- C) Organización de la información.
- M-7 Aprender a observar, a experimentar y a tomar datos de la realidad, para clasificarlos y cuantificarlos.

Relación de dichos objetivos con sus correspondientes de área y etapa

La selección descrita anteriormente cobraría mayor relevancia al relacionarlos con los objetivos para el área y para el ciclo. Tal relación puede observarse en el cuadro siguiente:

OBJETIVOS	OBJETIVOS DE ÁREA	OBJETIVOS DE NIVEL
M-1	1	b/e
M-2	1/7	b/e
M-3	2/4/7/8	c/d
M-4	1/4	b/e/h
M-6	1/4/7	b/e
M-7	6/7	c/d/i

3.4.3. Especificación del dominio

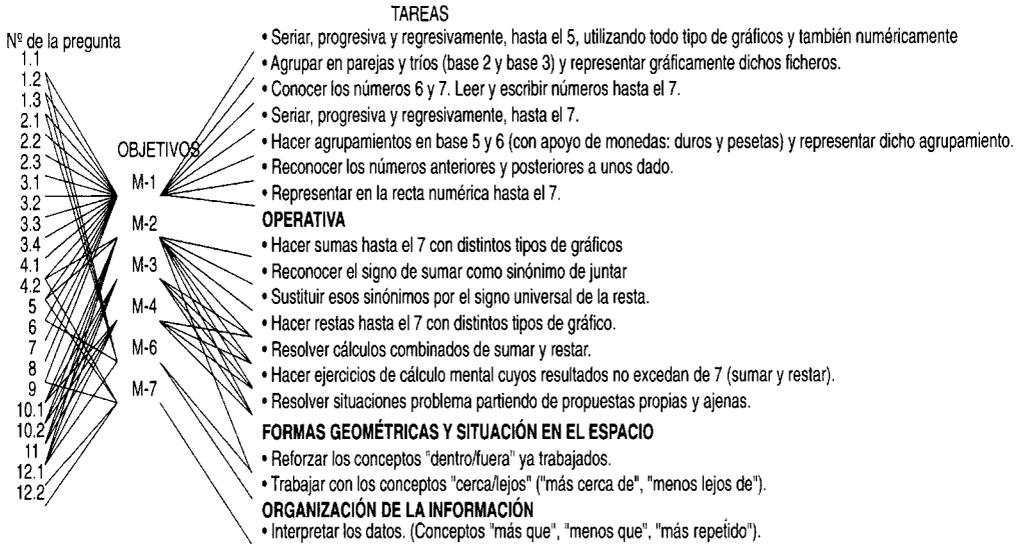
La producción de ítems para medir objetivos no puede realizarse eficazmente cuando los dominios de contenido no están definidos claramente y en consecuencia no es posible asegurar muestras representativas de ítems. Lo idóneo es elaborar **especificaciones de dominio** para clarificar el contenido que se pretende o las conductas especificadas para cada objetivo. Las especificaciones de dominio fueron acompañadas de normas de elaboración, de corrección, y de posibles distractores para la respuesta de cada ítems o tarea.

3.4.4. Primera redacción de ítems

Una vez establecidos los diferentes conocimientos y habilidades que deberían comprender cada objetivo, estos se transformaron en tareas, hipotetizando la relación directa entre cada pregunta o tarea y el objetivo que se pretende cubrir. El resultado de estas relaciones puede verse en el siguiente cuadro¹.

¹ El número de las preguntas corresponde a las diferentes tareas de la prueba criterial, no incluida por cuestión de espacio.

CORRESPONDENCIA ENTRE OBJETIVOS PROGRAMADOS, TAREAS REALIZADAS Y PREGUNTAS DE LA PRUEBA DE EVALUACIÓN



3.4.5. Validación de contenido

Coincidimos con Hambleton (1990) que dado el propósito de la prueba, evaluación interna en el aula con función formativa, lo más importante es demostrar su validez de contenido.

Consideramos, que determinar la validez para cada ítem implica, evaluar en cada uno, estos tres rasgos:

Coherencia o grado en que la pregunta formulada se ajusta al objetivo que dice medir.

Representatividad o grado en que la pregunta formulada es considerada como la más acertada de todas las posibles.

Calidad técnica o grado en que la pregunta expresa con claridad, sin pistas gramaticales falsas, la acción que se espera debe realizar el alumno.

Para cada una de estas categorías se dieron, a los jueces expertos, suficiente y preparado material para que ellos expresaran los juicios que cada uno de los ítems les merecían en relación con los rasgos citados.

En el análisis de resultados procedentes de esta validación se utilizó la formulación propuesta por Rovinelli y Hambleton (1977):

$$I_{jk} = \frac{N}{2N-2} (\mu_k - \mu)$$

donde:

- I_{jk} = índice de congruencia ítem-objetivos.
- N = nº de objetivos totales evaluados por los jueces.
- μ_k = media de valoración del ítem en el objetivo k.
- μ = media del ítem en todos los objetivos.

Se fijaron criterios de adecuación para cada una de las categorías analizadas:

Un criterio absoluto, determinado a priori con el valor de 0.75 (que se considera un índice bastante adecuado para cuatro jueces).

Un criterio relativo, calculado para cada categoría, que tendría el valor de la media de los índices de congruencia obtenidos en cada una de ellas. Los valores conseguidos fueron:

- Coherencia 0.72
- Representatividad 0.66
- Calidad Técnica 0.69

3.4.6. Segunda redacción de ítems, examen de campo y ensamblaje de la prueba

Tras la información resultante de la validación de contenido se pasó a una segunda revisión de ítems y posterior modificación y redacción de aquéllos considerados como posibles o aceptados, revisando el aspecto cuyo índice de congruencia obtenido no superaba el criterio relativo.

También nos pareció que los profesores debían remitirnos información sobre estándares que consideraban válidos para estimar superados los diferentes objetivos (uno para cada uno), así como criterios de corrección que estimaran oportunas y toda aquella información que consideraran relevante para la administración de la prueba.

Finalmente se consensuó con los profesores la longitud total y el número de preguntas idóneo para cada objetivo. Prepararemos la **redacción definitiva**, las **claves de puntuación** y el **punto de corte**.

Recibida toda esta información y modificada la prueba se preparó un formato definitivo, dando instrucciones precisas a los profesores sobre la flexibilidad en la administración, recordándoles que este tipo de pruebas deben de adaptarse a las características de los alumnos.

4. PARTICIPANTES

Han participado doce profesores (grupo de alumnos) de primero de Educación Primaria de diferentes escuelas de Granada y provincia, todas públicas. El número de alumnos ha sido 220, la edad media del profesorado ha sido 35 años y la de los niños 5 años y 9 meses, en el momento de comenzar la experiencia.

En estas doce clases, cuatro trabajaban los alumnos en grupos colaborativos, tres en grupos cooperativos y cinco en enseñanza individual. En uno de estos, se trabajaba en

pequeño grupo (generalmente de cuatro), pero al tener metas individuales y realizar las tareas individualmente, no ha sido considerado por nosotros grupo de aprendizaje cooperativo ni colaborativo.

5. ANÁLISIS DE DATOS

Con los datos obtenidos se han realizado diferentes análisis en función del objetivo que nos interesaba cubrir.

5.1. Análisis Factorial

Para establecer las dimensiones básicas sobre las que se fundamenta la actuación docente y para determinar la validez de la escala de observación recurrimos al Análisis Factorial. El procedimiento específico utilizado fue el método de componentes principales, siguiendo el criterio de Kaíser para la determinación del número de componentes (valores propios mayores de la unidad) utilizando rotación ortogonal (criterio Varimax).

El programa utilizado ha sido el BMDP.

Previamente a la realización del análisis factorial, comprobamos la idoneidad de la matriz para dicho análisis. El determinante de la matriz de correlaciones es muy bajo (.0000018), no obstante se realizó el test de esfericidad de Bartlett. El valor obtenido fue suficientemente alto (Chi cuadrado $X^2 = 630.13$) y supone una $p=.000$. Es decir rechazamos la hipótesis nula, lo que supone aceptar que la matriz de correlaciones no es una matriz de identidad y existen intercorrelaciones significativas dado que el valor hallado en la prueba de esfericidad es alto.

Todo esto nos permite considerar la matriz adecuada para proceder al análisis factorial.

Los coeficientes de correlación múltiple entre cada una de las variables y todas las demás son muy altas (por encima de .70), por lo que no eliminamos ninguna variable.

Los resultados revelaron la existencia de cuatro factores que explican más del 90% de la varianza total.

Componente I

El primer componente que aparece hace referencia a los procedimientos y estrategias que utiliza el profesor en el aula. Saturan todas las variables de la escala, aunque las de mayor peso, una vez sometido a rotación, son las que constituyen la categoría de **profesor mediador de aprendizajes significativos**.

La varianza explicada por este factor antes de la rotación es de 11.2514, lo que supone un 65.76% de la varianza total.

Componente II

El segundo componente puede ser denominado **clima de clase**. La varianza que

SORTED ROTATED FACTOR LOADINGS (PATTERN)

		FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4
INFORMAC	13	0.951	0.000	0.000	0.000
IDEAS	18	0.924	0.000	0.000	0.266
PLANIFIC	8	0.884	0.256	0.340	0.000
DIÁLOGO	17	0.878	0.000	0.353	0.000
REFLEXIO	15	0.819	0.527	0.000	0.000
CONEXIÓN	12	0.800	0.553	0.000	0.000
PROCESOS	14	0.752	0.324	0.353	0.000
ESTEREOT	4	0.625	0.000	0.290	0.334
RESPECTO	1	0.000	0.865	0.000	0.264
SEGURIDA	2	0.476	0.804	0.000	0.000
CLIMA	7	0.412	0.776	0.373	0.000
ANIMACIO	19	0.000	0.744	0.503	0.000
DEBATES	16	0.000	0.000	0.909	0.253
AUTORIDA	3	0.000	0.482	0.844	0.000
MULTIDIR	5	0.483	0.410	0.668	0.000
PROBLEMA	11	0.000	0.566	0.645	0.000
ALUMNOS	10	0.464	0.477	0.553	0.384
INFORMAL	6	0.000	0.283	0.000	0.937
PERSONAS	9	0.000	0.000	0.384	0.833
VP		6.614	4.454	3.768	2.275

explica es de 2.4513 elevándose a 80.08% el porcentaje de varianza acumulada. Los elementos que más saturan en este factor (1, 2, 7 y 19) hacen referencia a las relaciones del profesor con los alumnos y al clima de clase que se genera.

Componente III

El tercer componente hace referencia a la **organización del aprendizaje**. La varianza explicada es 1.8423. Esto supone el 90.85% de la varianza acumulada. Las variables que más saturan son el 16, 3, 5, 10 y 11. Estos elementos están referidos a la forma de organizar la participación del alumno en la adquisición de los conocimientos.

Componente IV

El cuarto componente se refiere a la **comunicación informal** que se establece en el aula. La varianza explicada es 1.5655. Con este factor explicamos el 100% de la

varianza acumulada. Las variables con mayores pesos son la 6 y la 9 que hacen referencia a la comunicación sobre contenidos no curriculares y la aceptación de personas ajenas al aula.

5.2. Análisis de Covarianza

En segundo lugar procedimos a comprobar la hipótesis de investigación, que planteada en términos de hipótesis nula es:

«No existen diferencias significativas en el nivel de logro de los objetivos de Primaria en Matemáticas en función de la metodología de aprendizaje del alumno».

Para comprobar esta hipótesis hemos seguido el mismo diseño estadístico que planteó Robert Linn (1986: 102-103) para la evaluación de los programa **Head Start**. En este diseño actúan de covariantes variables referidas a los alumnos y variables referidas a profesores. En nuestro caso el diseño seguido ha sido:

Variable Independiente	Metodología con tres niveles
1. Individual	
2. Cooperativa	
3. Colaborativa	
Variable Dependiente	Objetivos de Matemáticas
M1, M2, M3, M4, M5, M6	

Covariantes:

Referidas al profesor:

- a) Modelo adulto significativo
- b) Grupo humano cohesionado
- c) Mediador de aprendizajes significativos

Referidas al alumno:

- a) Aptitud general verbal
- b) Aptitud general no verbal

Con los datos obtenidos se realizaron seis ANCOVAS con el paquete SYSTAT, para Macintosh.

El nivel de significación establecido para la comprobación de hipótesis fue el .05.

Resultados del ANCOVA

Exponemos sólo los resultados del ANCOVA, aunque previamente se han probado diferentes modelos estadísticos sencillos y progresivamente hemos ido introduciendo mayor control hasta llegar al modelo propuesto. Comenzamos con un ANOVA

entre los tres grupos de Metodología, encontrando diferencias significativas entre los tres niveles de la variable independiente.

Antes de aplicar el ANOVA se comprobaron los supuestos de Normalidad, Homocedasticidad e Independencia. Aunque la Normalidad sea irrelevante para los resultados del ANOVA (Tejedor, 1988: 281), no obstante se realizaron los índices de Asimetría y Curtosis puesto que después la homocedasticidad se comprobó con el contraste de Bartlett fuertemente afectado por la no-normalidad.

La Asimetría fue próxima a cero y la curtosis 2,83.

La prueba de Bartlett nos permitió aceptar la homogeneidad de las varianzas al nivel de significación del.05.

Posteriormente realizamos un ANCOVA con una sola covariante, aptitud verbal del alumno. El modelo resultó válido ($F < .01$) y las diferencias en el nivel de logro de cada uno de los objetivos varió sensiblemente. Por último y considerando que la variable de actuación docente podía influir igualmente en la variable dependiente, realizamos los ANCOVAs que exponemos a continuación.

Primer Objetivo: Conocer los principales sistemas de numeración.

DEP VAR: M1 N: 193 MULTIPLE R: 0.648
SQUARED MULTIPLE R: 0.419

ANALYSIS OF VARIANCE					
SOURCE	SUM-OF-SQU	DF	MEAN-SQUARE	F-RATIO	P
METODO	7.773	2	3.887	0.533	0.588
A	26.363	1	26.363	3.617	0.059
B	0.191	1	0.191	0.026	0.872
C	67.668	1	67.668	9.284	0.003
AGV	121.781	1	121.781	16.709	0.000
AGNV	157.241	1	157.241	21.574	0.000
ERROR	1348.373	185	7.289		

En el nivel de logro de este primer objetivo, no existen diferencias significativas en función de la metodología de trabajo del alumno.

El modelo es válido puesto que tanto una de las variables referidas al profesor (crea un grupo humano cohesionado) como las aptitudes de los alumnos, influyen significativamente en la variable dependiente

Segundo Objetivo: Conocer y manejar operaciones elementales de cálculo: sumas y restas.

DEP VAR: M2N: 193 MULTIPLE R: 0.633
SQUARED MULTIPLE R: 0.400

ANALYSIS OF VARIANCE					
SOURCE	SUM-OF-SQU	DF	MEAN-SQUARE	F-RATIO	P
METODO	6.962	2	3.481	1.694	0.187
A	18.848	1	18.848	9.170	0.003
B	2.486	1	2.486	1.210	0.273
C	37.535	1	37.535	18.261	0.000
AGV	27.623	1	27.623	13.439	0.000
AGNV	36.113	1	36.113	17.569	0.000
ERROR	380.260	185	2.055		

En este segundo objetivo tampoco encontramos diferencias significativas en función de la Metodología. Vuelven a aparecer las aptitudes de los alumnos y las variables del profesor, influyendo significativamente en el nivel de logro de los objetivos.

Tercer Objetivo: Utilizar la operativa para resolver situaciones problema de la vida cotidiana.

DEP VAR: M3 N: 193 MULTIPLE R: 0.611
SQUARED MULTIPLE R: 0.373

ANALYSIS OF VARIANCE					
SOURCE	SUM-OF-SQU	DF	MEAN-SQUARE	F-RATIO	P
METODO	7.476	2	3.738	3.617	0.029
A	4.652	1	4.652	4.502	0.035
B	3.730	1	3.730	3.610	0.059
C	16.873	1	16.873	16.329	0.000
AGV	9.044	1	9.044	8.752	0.003
AGNV	16.796	1	16.796	16.255	0.000
ERROR	191.158	185	1.033		

El nivel de logro del tercer objetivo si es significativamente diferente en función de la metodología. Haciendo la diferencia entre las medias ajustadas encontramos que son los alumnos que han trabajado en grupos cooperativos los que mejor han superado este objetivo. Es igualmente significativa la influencia que ejercen las variables de aptitudes de los alumnos y las referidas al profesor (denominadas globalmente «ayudas» sobre la consecución de este objetivo.

Cuarto Objetivo: Efectuar operaciones sencillas de cálculo mental, mostrando confianza en las propias capacidades.

DEP VAR: M4 N: 193 MULTIPLE R: 0.611

SQUARED MULTIPLE R: 0.373

ANALYSIS OF VARIANCE					
SOURCE	SUM-OF-SQU	DF	MEAN-SQUARE	F-RATIO	P
METODO	7.476	2	3.738	3.617	0.029
A	4.652	1	4.652	4.502	0.035
B	3.730	1	3.730	3.610	0.059
C	16.873	1	16.873	16.329	0.000
AGV	9.044	1	9.044	8.752	0.003
AGNV	16.796	1	16.796	16.255	0.000
ERROR	191.158	185	1.033		

Los resultados son iguales que en el tercer objetivo, puesto que se han dado respuesta a ambos con idénticos ítems.

Quinto objetivo: Mostrar interés en la búsqueda de soluciones, no fue incluido en la prueba, por la dificultad que a juicio de los profesores suponía el conocimiento de este objetivo a través de una prueba escrita.

Sexto Objetivo: Lograr un progresivo dominio de los conceptos espaciales.

DEP VAR: M6 N: 193 MULTIPLE R: 0.548

SQUARED MULTIPLE R: 0.301

ANALYSIS OF VARIANCE					
SOURCE	SUM-OF-SQU	DF	MEAN-SQUARE	F-RATIO	P
METODO	2.925	2	1.463	1.110	0.332
A	4.061	1	4.061	3.081	0.081
B	0.441	1	0.441	0.335	0.564
C	11.920	1	11.920	9.043	0.003
AGV	19.187	1	19.187	14.556	0.000
AGNV	6.983	1	6.983	5.298	0.022
ERROR	243.854	185	1.318		

En la consecución del sexto objetivo, tampoco existen diferencias significativas en función de la metodología de trabajo del alumno. Vuelven a aparecer las variables de aptitudes de los alumnos y las referidas al profesor con un fuerte peso sobre la variable dependiente.

Séptimo Objetivo: Aprender a observar, a experimentar y a tomar datos de la realidad para clasificarlos y cuantificarlos.

DEP VAR: M7 N: 193 MULTIPLE R: 0.592
SQUARED MULTIPLE R: 0.351

ANALYSIS OF VARIANCE					
SOURCE	SUM-OF-SQU	DF	MEAN-SQUARE	F-RATIO	P
METODO	1.540	2	0.770	0.465	0.629
A	12.326	1	12.326	7.448	0.007
B	0.008	1	0.008	0.005	0.944
C	15.987	1	15.987	9.660	0.002
AGV	18.716	1	18.716	11.309	0.001
AGNV	24.194	1	24.194	14.619	0.000
ERROR	306.166	185	1.655		

Tampoco son significativas las diferencias en el nivel de logro del objetivo que se pretende conseguir, en función de la metodología, siendo igualmente significativa la influencia de las restantes variables sobre la variable dependiente.

Esto nos pone de manifiesto la importancia que tienen las mediaciones, ayudas, que el profesor utiliza en el proceso de enseñanza-aprendizaje. La influencia del trabajo individual o cooperativo queda encubierta ante la importancia que esta variable ocupa en el aula, junto con las aptitudes verbales y numéricas de los alumnos.

6. CONCLUSIONES

De los resultados obtenidos podemos concluir:

1. En las tareas sobre sistemas de numeración y operaciones de cálculo no existieron diferencias significativas en función del tipo de metodología de trabajo. Sin embargo, sí influyeron significativamente en el nivel de logro conseguido por los alumnos las ayudas facilitadas por el profesor, sobre todo: el profesor ayuda a planificar e interviene en el trabajo del grupo.

2. Los profesores intervienen y proporcionan más ayudas cuando los alumnos trabajan cooperativamente resolviendo problemas de la vida cotidiana.

Salvo en este caso, en el que la variable metodología de trabajo se presenta como altamente significativa, en el resto de las tareas, las variables tipo de ayuda facilitada

por el profesor y aptitudes de los alumnos, han sido las que mayor influencia han tenido.

Los **rasgos de actuación docente o ayudas** que mayormente utilizan los profesores observados, por orden de prioridad son:

- 1º Observa el proceso que siguen los alumnos y se detienen en aquellos aspectos que parecen no estar claros.
- 2º Planifican e interviene en el trabajo intelectual.
- 3º Ofrecen información relevante.
- 4º Establecen conexiones entre el conocimiento nuevo y el que el niño ya tenía.

Con menor frecuencia, otras ayudas que aparecen son:

- Establece relaciones entre los alumnos.
- Anima a intervenir.
- Participación solidaria en el grupo.
- Establece una red de comunicación multidireccional.
- Crea situaciones que permiten al alumno reflexionar sobre su propio proceso de aprendizaje.

Donde mayores diferencias de actuación han existido ha sido en las categorías A y B. (Sirve de modelo adulto significativo y crea un grupo humano cohesionado). Hay ayudas que son dominantes a lo largo de todo el proceso de observación en el 80% de los profesores observados como son: Planifica e interviene en el trabajo del grupo (de la categoría B) y consigue el respeto y la confianza del niño (de la categoría A); y otros rasgos: favorece la comunicación informal (categoría B) o manifiesta opiniones que sirven para combatir estereotipos sociales, (categoría A), que no aparecen nunca.

En general, podemos concluir que el papel mediador del profesor en el aula es importantísimo en la construcción del conocimiento del alumno. En esta construcción, las aportaciones de alumnos y profesor pueden ser asimétricas (Coll, 1992) pero desde luego lo que no se puede es, desde posiciones constructivistas, olvidar el importante papel que juega el profesor en esta construcción, orientando y mediando durante todo el proceso.

BIBLIOGRAFÍA

- ALLEN, V.L. y FELDMAN, R.S. (1973). Learning through tutoring: Low-achieving children as tutors. *Journal of Experimental Education*. 42, 1-5.
- ARONSON, E. y OSHEROW. (1980). Cooperation, prosocial behavior, and academic performance. En L. Bickman (Ed.). *Applied social psychology annual*. Vol. 1, pp. 163-196. Sage. Beverly Hills.
- ARONSON, E.; BLANEY, N.; STEPHAN, C.; SIKES, J. y SNAPP, M. (1978). *The Jigsaw Classroom*. Sage. Beverly Hills, Cal.

- ARONSON, E. y BRIDGEMAN, D. (1981). Jigsaw groups and the desegregated classroom. In pursuit of commongoes. In E.M. Hetherington y R.D. Parke (Eds.): *Contemporary readings in child psychology*. (2nd ed. pp. 339-345). McGraw Hill. New York.
- BENNET, N. y DUNNE, E. (1991). The nature and quality of talk in co-operative classroom groups. *Learning and Instruction*, vol. 1, 103-118.
- BENNETT, N.; DESFORGES, C.; COCKBURN, A. y WILHINSON, B. (1984). *The quality of pupil learning experiences*. Erlbaum. London.
- BENNETT, N. y KELI, J. (1989). *A good start? Four year olds in infant schools*. Blackwell. Oxford.
- BIJOU, S.W.; PETERSON, R.F.; HARRIS, F.R.; ALLEN, K.E. y JOHNSON, M.S. (1969). Methodology for experimental studies of young children in natural settings. *The psychological Record*, 19, 177-210.
- BRUNER, J. (1986). *Actual minds, possible worlds*. Harvard University Press. Cambridge, MA.
- BUENDÍA, L. y FERNÁNDEZ, A. (1991). Una experiencia exploratoria en cognición numérica en escolares de 3º de E.G.B. *Bordón*. 41 (1).
- BUENDÍA, L. y SALMERÓN, H. (1994). Construcción de pruebas criterios de aula. *Revista de Investigación Educativa*. 23, 405-411.
- BUENDÍA, L.; RUIZ, J. y FERNÁNDEZ, A. (1995): Estudios evaluativos en diferentes contextos. *Revista de Investigación Educativa*. 26, 159-185.
- BUENDÍA, L. (1993). Los mecanismos de influencia educativa en el logro de los objetivos de Educación Primaria. Investigación realizada para el segundo ejercicio de cátedra de Universidad. Universidad de Granada.
- BUENDÍA, L. (1987). Estilo cognitivo y Lectura. *Cuestiones Pedagógicas*. 4, 23-31.
- BUENDÍA, L. (1985). Progreso académico en función del profesor. *Revista de Investigación Educativa*. 3 (5), 25-33.
- BUENDÍA, L. (1993). Diseño de una evaluación criterial de aula en el área de Lenguaje y de Conocimiento del Medio. Actas de: *Jornades sobre avaluació a l'aula*. ICE. Universidad de Barcelona. 181-187 y 244-250.
- COLL, C. y otros (1992). Actividad conjunta y habla: una aproximación al estudio de los mecanismos de influencia educativa. *Infancia y Aprendizaje*. 59-60, 189-232.
- DISEÑOS CURRICULARES DE LA REFORMA (1989). *Diseño curricular de educación primaria*. Junta de Andalucía. Consejería de Educación y Ciencia. Sevilla.
- DOISE, W.; MUGNY, G. y PERRET-CLERMONT, A.N. (1975). Social interaction and the development of cognitive operations. *European Journal of Social Psychology*. 5, 367-383.
- DUNNE, E. y BENNET, N. (1990). *Talking and learning in groups*. Macmillan. London.
- FORMAN, E.A. y CAZDEN, C.B. (1984). Perspectivas vygotskianas en la educación: el valor cognitivo de la interacción entre iguales. *Infancia y Aprendizaje*. 27/28, 139-157.
- FULLANA, J. (1996). La investigación sobre variables relevantes para la prevención del fracaso escolar. *Revista de Investigación Educativa*. 14, número 1. 63-90.
- GALTON, M.; SIMON, B. y CROLL, P. (1980). *Inside the Primary Classroom*. Routledge and Kegan Paul. Boston.

- GELERMAN, S.L. (1987). An observational study of small-group instruction in fourth-grade mathematics classrooms. *Elementary School Journal*. 88, 3-28.
- GOOD, T.; GROUWS, D. y MASON, D. (1990). Teachers' beliefs about small-group instruction in elementary school mathematics. *Journal for Research in Mathematics Education*, 21 (1), 2-15.
- GOODLAD, J.I. (1984). *A place called scholl*. McGraw-Hill. New York.
- GRAYBEAL, S.S. y STODOLSKY, S.S. (1985). Peer work groups in elementary schools. *American Journal of Educational*. 93, 409-428.
- GRAYBEAL, S.S. y STODOLSKY, S.S. (1985). Peer work groups in elementary schools. *American Journal of Educational*. 93, 409-428.
- HAMBLETON, R. (1990). Criterion referenced testing. Methods and practices. En Gutkin, T. y Reynolds, C. (Eds.). *The handbook of School Psychology*. Wiley. New York.
- HARRIS, F. y LAHEY, B.B. (1978). A method for combining occurrence and nonoccurrence interobserver agreement scores. *Journal of Applied Behavior Analysis*. 11, 523-527.
- HARTMANN, D.P. (1979). Inter e intra-observer agreement as a function of explicit behavior definitions in direct observation: A critique. *Behavioural Analysis and Modification*, 3 (4), 229-233.
- HOPKINS, B. y HERMANN, J. (1977). Evaluating interobserver reliability of interval data. *Journal of Applied Behavior Analysis*, 10, 121-126.
- JOHNSON, D.W. and JOHNSON, R.T. (1975). *Learning together and alone*. Prentice-Hall. Englewood Cliffs, NJ.
- KRIPPENDORF, K. (1980). *Contend analysis: An introduction to its methodology*. Sage. Beverly Hills, Ca.
- LOCKHEED, M.E. y HARRIS, A.M. (1984). Cross-sex collaborative learning in elementary classrooms. *American Educational Research Journal*. 21, 275-294.
- MYERS, D.G. y LAMM, H. (1976). The group polarization phenomenon. *Psychological Bulletin*. 83, 602-607.
- SHARAN, S. (1980). Cooperative learning in small groups: Recent methods and effects on achievement, attitudes and ethnic relations. *Review of Educational Research*. 50, 241-271.
- SLAVIN, R. (1989). Cooperative learning and student achievement. In R. Slavin (Ed.): *School an Classroom Organization*, (pp. 129-156). Erlbaum Hillsdale, NJ.
- TEJEDOR, F.J. (1988). *El análisis de varianza en Psicología y Educación*. Madrid. Anaya.
- VYGOTSKY, L.S. (1973). Aprendizaje y desarrollo intelectual en la edad escolar. En Luria, Leontiev, Vygotsky y otros (compilación). *Psicología y Pedagogía*. Akal. Madrid.
- WITTRUCK, M.C. (1974). Learning as a generative process. *Educational Psychologist*. 11, 87-95.
- YUSTE, C. (1985). *Batería de aptitudes diferenciales y generales*. C.E.P.E. Madrid.

EVALUACIÓN DE LA RESOLUCIÓN DE PROBLEMAS ARITMÉTICOS EN PRIMARIA¹

Enrique Castro; Luis Rico; José Gutiérrez; Encarnación Castro; Isidoro Segovia; Nicolás Morcillo;
Francisco Fernández; Evaristo González y Antonio Tortosa
Universidad de Granada

RESUMEN

Se analiza la dificultad de los niños de 4º, 5º y 6º de Primaria (9, 10 y 11 años) para seleccionar las operaciones adecuadas cuando resuelven problemas verbales de estructura aditiva de dos etapas. El objetivo es poner de manifiesto la competencia que tienen los niños en la resolución de estos problemas aritméticos y en qué medida se ve afectada por el tipo de relación (aumento o disminución) en la primera y segunda posición. El diseño es de tipo factorial mixto con dos variables intersujeto (colegio y curso) y una variable intrasujeto (relación de aumento o disminución en primera posición) con medidas repetidas en esta última variable. Las conclusiones obtenidas son: a) la capacidad de resolución de problemas aritméticos de los alumnos se desarrolla al pasar de 4º a 6º. b) La variable «tipo de relación» influye en la dificultad de elegir el proceso adecuado de solución.

Enrique Castro. Profesor Titular de Universidad. Departamento de Didáctica de la Matemática. Universidad de Granada. E-mail: ecastro@platon.ugr.es

¹ Este trabajo ha sido parcialmente financiado con una ayuda a la Investigación para Niveles no universitarios (BOJA, nº 131, 19 de diciembre de 1992), concedida por la Consejería de Educación y Ciencia de la Junta de Andalucía al Proyecto denominado *Diagnóstico de Procedimientos y Evaluación de Destrezas Terminales para la Resolución de Problemas Aritméticos en el Tercer Ciclo de la Educación Primaria Obligatoria*.

ABSTRACT

We analyze the difficulty of Primary School children of 4^o, 5^o and 6^o levels (9, 10 and 11 years old), into choice of the proper operations when they try to solve two-steps arithmetical word problems with additive structure. Our goal is to set up the ability of the pupils to solve such kind of arithmetic problem and to value the way in that the kind of relationship (increase or decrease) both in first or second position affects this ability. We have used a repeated measure design with two between-subjects factors (teaching institution and course) and one within-subjects factor (relationship in first and second position). The obtained conclusions are: a) the ability of the pupils in arithmetic problem solving is developed upon going from 4^o to 6^o. b) The variable «type of relationship» influences the difficulty of choosing the adequate process of solution.

INTRODUCCIÓN

El aprendizaje de la aritmética en los niveles escolares obligatorios está estrechamente ligado a las tareas de resolución de problemas. Los problemas aritméticos de enunciado verbal ocupan parte importante del tiempo reservado a matemáticas en los calendarios escolares, por otra parte los alumnos, a lo largo de su vida, encuentran multitud de problemas que deben resolver. Unas veces los problemas se plantean como estrategia de refuerzo y aplicación de aspectos conceptuales, mientras que otras veces se plantean como una tarea rutinaria para rellenar el horario de matemáticas. En cierto modo son una alternativa a la mecánica de *hacer cuentas sin un contexto explícito de referencia*.

Independientemente del éxito obtenido por los escolares en las tareas de resolución de problemas aritméticos, el objetivo fundamental para su tratamiento curricular se basa en la creencia de que la propuesta sistemática y reiterada de problemas a los alumnos va a producir una mejora progresiva en sus capacidades de resolución de problemas con el paso del tiempo. Esta creencia ha llevado, en ocasiones, a proponer problemas difíciles, con dos o tres operaciones, a alumnos en edades muy tempranas sin la necesaria madurez y desarrollo cognitivo.

Los investigadores en educación matemática han reconocido este ámbito de las matemáticas escolares como un campo fértil, digno de ser explorado más sistemáticamente en orden a comprender cómo los alumnos resuelven los problemas, las dificultades que se plantean, las variables relevantes en los procesos de resolución, los tipos diferentes de problemas que existen y los niveles educativos en que deben ser abordados.

Hoy día existe un campo de estudios e investigaciones sobre resolución de problemas aritméticos escolares con una producción considerable de literatura especializada, que estudia la influencia de variables tales como la estructura semántica, la posición de la incógnita o el tipo de relaciones de aumento o disminución entre las cantidades implicadas en el problema.

Una revisión actualizada (Castro, 1995; Castro, Rico y Gil, 1992) de las diferentes corrientes investigadoras centradas en el estudio global o parcial de las tareas de

resolución de problemas aritméticos permite distinguir tres grandes enfoques: a) *Estudios correlacionales centrados en factores lingüísticos*; b) *estudios basados en variables estructurales*, y c) *estudios de tipo semántico*.

Los estudios de tipo semántico sobre problemas aditivos han estado dedicados principalmente a aquellos en cuya solución interviene una sola operación, problemas de una etapa, y constituyen la clase más sencilla de problemas aditivos. Las categorías semánticas para problemas aditivos de una etapa, que representan estructuras alternativas de información cuantitativa, fueron establecidas inicialmente por Heller y Greeno (1979) y, con sucesivos perfeccionamientos, han llegado hasta nuestros días y se mantienen como criterios válidos en las agendas de investigación sobre Educación Matemática. Fuson (1992) recoge 22 problemas aditivos estructuralmente diferentes, para cuya elaboración tiene en cuenta el cruce entre: a) cuatro estructuras semánticas alternativas, que se denominan Combinación, Cambio, Comparación e Igualación; b) dos tipos de relaciones —aumento o disminución— para las tres últimas categorías, o la disyuntiva entre una relación estática o dinámica para la primera estructura; c) tres posibilidades para el dato desconocido en la estructura de relaciones que establece el enunciado del problema (solamente dos en la estructura de Combinación).

El trabajo se sitúa dentro de este marco, siendo uno de los objetivos principales en este estudio determinar cuáles son las variables estructurales de los problemas aritméticos y qué influencia pueden ejercer sobre la dificultad en la resolución de los mismos. Se estudian problemas aritméticos aditivos, clasificados según las cuatro categorías de estructura semántica citadas. Se centra en el análisis de los problemas aritméticos aditivos de dos etapas o de dos pasos.

Problemas de dos pasos

Un problema aritmético de dos pasos es el que requiere dos operaciones para obtener la solución numérica del problema. Por ejemplo:

Problema. En un autocar había 19 pasajeros. En la primera parada bajan 8 pasajeros y suben 5. ¿Cuántos pasajeros hay ahora en el autocar?

Realizando las combinaciones de los cuatro procesos básicos (adición, sustracción, multiplicación y división) de dos en dos, obtenemos las dieciséis parejas de operaciones:

(+, +), (+, -), (+, x), (+, :), (-, +), (-, -), (-, x), (-, :),
(x, +), (x, -), (x, x), (x, :), (:, +), (:, -), (:, x), (:, :)

que darán lugar a los enunciados de otros tantos problemas de dos pasos o etapas cuya solución requiera de las operaciones básicas. No hay muchos trabajos previos de investigación sobre los problemas compuestos de varias etapas. Vamos a citar aquellos que nos parecen más interesantes para nuestro estudio.

Partiendo del trabajo teórico de Durell (1940), Gray & Young utilizan la idea de la categorización de los problemas según las operaciones que intervienen en su solución para comparar niveles de dificultad entre problemas de dos pasos.

Berglaund-Gray (1938, 1939), Berglaund-Gray & Young (1940), y Young & McIsaac (1941) investigan el efecto de invertir la secuencia de operaciones. Concluyen que la sustracción y la división fueron más difíciles como primeros procesos y la adición y la multiplicación fueron más difíciles como segundos procesos.

Suppes, Loftus y Jerman (1969) obtienen que «el número mínimo de operaciones diferentes requeridas para obtener la solución correcta» tiene influencia significativa sobre la dificultad de resolución de un problema aritmético; pero no así el número de pasos requerido para obtener la solución. En el trabajo posterior de Jerman y Rees (1972), también se concluye que las operaciones matemáticas tienen efecto significativo para predecir el orden de dificultad de los problemas verbales.

Sherard (1974) halla, para problemas de dos pasos, que los problemas que implican a la multiplicación y a la división son los más difíciles, siguiendo en orden descendente de dificultad los de sustracción y multiplicación, sustracción y división, adición y división, adición y multiplicación, y adición y sustracción.

El interés por el estudio de los problemas aritméticos de dos pasos (o etapas) ha renacido recientemente (Shallin, 1985; Nesher, 1991; Nesher y Herskovitz, 1994). Los trabajos se centran en el estudio de los esquemas subyacentes a los problemas de dos etapas. Estos autores consideran que los esquemas de los problemas de dos etapas son esquemas compuestos por dos esquemas simples: el aditivo y el multiplicativo. Parten del hecho de que la definición de problema compuesto no se basa en las cuatro operaciones aritméticas sino en las esquemas simples constituidos por una relación entre tres cantidades. En el esquema aditivo las tres componentes son el todo y sus dos partes, mientras que en el esquema multiplicativo las componentes son el producto y sus dos factores. Los autores se basan en el trabajo previo de Shallin y Bee (1985) para afirmar que sólo hay tres esquemas básicos de problemas de dos etapas:

Esquema (A) (*Jerárquico*). El todo de un esquema es una parte del otro esquema.

Esquema (B) (*Compartir el todo*). Los dos esquemas comparten un todo.

Esquema (C) (*Compartir una parte*). Los dos esquemas comparten una parte.

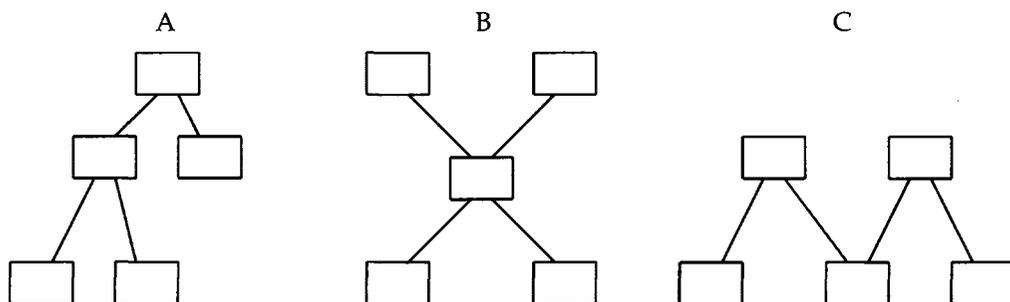


Ilustración 1

Esquemas básicos de problemas de dos etapas (Nesher y herskovitz, 1994).

Una de las cuestiones que investigan es la influencia que tienen sobre el índice de dificultad de los problemas de dos etapas los tres esquemas y las ocho posibles combinaciones de dos operaciones que se obtienen eligiendo una operación de estructura aditiva + / - y otra de estructura multiplicativa \times / : a saber:

$$(+, \times), (\times, +), (+, :), (:, +), (-, \times), (\times, -), (-, :), (:, -)$$

Para una muestra de escolares de 3º, 4º, 5º y 6º de Israel obtienen que ambas variables tienen efecto significativo sobre el índice de dificultad de estos problemas.

Con respecto a los esquemas hallan que el esquema «jerárquico» es el más fácil, le sigue en dificultad el esquema «compartir el todo» y, por último, el esquema «compartir una parte».

Categorización de los problemas en este estudio

En nuestra investigación estudiamos problemas de dos etapas desde el punto de vista estructural basándonos en las categorías semánticas de los mismos. Problemas aritméticos aditivos de dos etapas son los problemas cuyas soluciones implican solamente sumas y restas y, en todos los casos, es necesario realizar dos operaciones —distintas o repetidas— para obtener la solución; el estudio de los problemas aritméticos aditivos de dos etapas constituye el objeto principal de esta investigación. Uno de los enfoques de investigación más productivos en los estudios sobre problemas aritméticos de estructura aditiva de una etapa es el enfoque estructural basado en las categorías semánticas de los problemas (Castro y otros, 1992). Las cuatro categorías semánticas determinadas para estos problemas constituyen uno de los cimientos de nuestra categorización teórica; criterio que se complementa con otras dos variables: el tipo de relación y la posición de la incógnita.

a) *Estructura semántica de la primera y segunda etapa.* Partimos del supuesto teórico de que, en cada etapa, está presente una de estas estructuras semánticas y que, por tanto, al tratarse de problemas de dos etapas hemos de tener en cuenta la categoría semántica que se utiliza en cada una de ellas. Dicho de otra manera, en el enunciado de los problemas aritméticos de dos etapas, cada etapa del problema corresponde a una relación entre tres datos, uno de los cuales es compartido y está latente en el problema. Cada una de estas relaciones por separado las encontramos en los problemas aritméticos verbales de una etapa. En consonancia con esto, en un problema verbal de dos etapas la estructura semántica se encuentra tanto en la primera relación como en la segunda. Utilizamos, por tanto, las categorías semánticas en la primera relación del problema y las categorías semánticas en la segunda relación. Las categorías semánticas empleadas son las denominadas:

- **cambio**, se refiere a los problemas en los que se produce algún evento o transformación que cambia el valor de una cantidad inicial; la codificamos como Ca.
- **combinación**, problemas basados en la relación existente entre un conjunto total y una partición del mismo en dos subconjuntos; la codificamos como Co.

- **comparación**, son problemas que implican una relación comparativa entre dos cantidades; la designamos abreviadamente como Cp.
- **igualación**, son aquellos problemas en los que se plantea una acción para lograr que una cantidad sea igual a otra; abreviadamente lo designamos como Ig.

Si atendemos a las posibilidades que ofrecen estas cuatro estructuras en los problemas de dos etapas, encontramos las 16 opciones que están recogidas en la Tabla 1.

TABLA 1
COMBINACIONES SEMÁNTICAS POSIBLES EN PROBLEMAS DE DOS ETAPAS

	Ca	Co	Cp	Ig
Ca	(Ca, Ca)	(Ca, Co)	(Ca, Cp)	(Ca, Ig)
Co	(Co, Ca)	(Co, Ca)	(Co, Cp)	(Co, Ig)
Cp	(Cp, Ca)	(Cp, Co)	(Cp, Cp)	(Cp, Ig)
Ig	(Ig, Ca)	(Ig, Co)	(Ig, Cp)	(Ig, Ig)

b) *Relaciones de Aumento/Disminución*. La segunda variable hace referencia a los tipos de relaciones de comparación posibles en los problemas aditivos de dos etapas. Cuando enunciamos un problema verbal, la relación entre cada pareja de datos puede ser de aumento o disminución (Carpenter y Moser, 1983). En este caso cada una de las etapas incluye una relación de aumento o disminución, por lo que se presentan cuatro posibilidades para la variable relaciones, recogidas en la Tabla 2.

TABLA 2
COMBINACIONES SEGÚN RELACIONES POSIBLES EN PROBLEMAS DE DOS ETAPAS

	Aumento (A)	Disminución (D)
Aumento (A)	Aumento-Aumento (A,A)	Aumento-Disminución (A,D)
Disminución (D)	Disminución-Aumento (D,A)	Disminución-Disminución (D,D)

c) *Cantidad desconocida*. Una tercera caracterización de los problemas de dos etapas se refiere a cuál es la cantidad desconocida entre las posibles que intervienen en la relación global subyacente en el problema. En nuestro estudio sobre problemas aditivos de dos etapas hemos decidido trabajar exclusivamente sobre los problemas en los que el resultado de una operación entra como dato para la otra operación; con mayor precisión, consideramos problemas en cuyo enunciado aparecen en primer lugar dos datos con los que hay que operar para obtener un tercer número, este primer resultado hay que operarlo a su vez con el tercer dato del enunciado para alcanzar la solución. Esquemáticamente:

- * datos ordenados del problema: a, b, c
- * orden de operaciones para alcanzar la solución:

$$a * b \rightarrow d$$

$$d * c \rightarrow \text{solución}$$

En definitiva, para el estudio de los problemas aritméticos de estructura aditiva de dos etapas hemos delimitado tres características clave: a) la **estructura semántica**, con dieciséis posibilidades; b) las **relaciones de comparación**, con cuatro posibilidades; y c) la **estructura ordenada de las operaciones**, con la «cantidad desconocida» siempre al final. Hemos construimos así un universo de 64 problemas aritméticos de dos etapas, en base a las características anteriores, manteniendo controladas otra serie de características en cada uno de los enunciados, evitando así una posible interferencia debida a estos factores.

Objetivos

En la resolución de un problema han sido identificadas dos fases generales: la comprensión del problema y la solución del problema (Kintsch y Greeno, 1985; Mayer, 1986a, 1986b; Newell y Simon, 1974; Riley, Greeno y Heller, 1983).

Varios estudios sobre resolución de problemas aritméticos han mostrado que la mayoría de los errores que cometen los estudiantes en problemas verbales se deben a falta de comprensión de la estructura del problema más que a errores de cálculo. Los estudiantes pueden ser capaces de realizar determinados cálculos pero no ser capaces de resolver problemas verbales en los que para obtener la solución sólo se requiere de esos cálculos (Carpenter, Corbitt, Kepner, Lindquist y Reys, 1980; De Corte, Verschaffel y De Win, 1985; Dellarosa Cummins, Kintsch, Reusser y Weimer, 1988).

Esta discrepancia en el índice de dificultad de un problema según el formato numérico o verbal sugiere que la dificultad que plantea un problema a los niños no está determinada sólo por las destrezas de cálculo implicadas sino que hay otros aspectos que contribuyen a su dificultad, entre ellos, aspectos relacionados con la fase de comprensión. Por ello, nuestro trabajo tiene como objetivo evaluar la dificultad de resolución de problemas aritméticos verbales de dos pasos, con estructura aditiva, centrándonos en la fase de comprensión.

Las cuestiones que nos planteamos al respecto son:

- 1) *¿Hay diferencias en el rendimiento de los alumnos en la resolución de problemas aritméticos verbales de dos etapas según el colegio y el curso escolar en el que se hallen?*
- 2) *¿Tienen los distintos problemas aritméticos verbales compuestos de dos etapas el mismo índice de dificultad de comprensión o ésta depende de la variable de tarea combinación de relaciones de aumento y de disminución?*

MÉTODO

Muestra

Las pruebas se han aplicado en 6 colegios diferentes de Granada y provincia. De los centros que participaron en la experiencia cinco pertenecen a Granada capital y uno a un pueblo de la provincia. De los colegios de la capital dos están situados en la Zona centro de Granada y los otros tres están ubicados en distintos barrios de la ciudad.

Con relación a su carácter administrativo, tres tienen establecido concierto con la Administración Educativa y los tres restantes son públicos.

En cada uno de los seis colegios se han elegido simultáneamente los cursos 4º, 5º y 6º de Primaria. De cada colegio, las pruebas se han aplicado a un grupo completo de alumnos por cada uno de los cursos. En total, el número de grupos que componen la muestra son 18.

Diseño e instrumento

Hemos utilizado un diseño factorial mixto en el que han intervenido las siguientes variables:

a) Variables independientes

Las variables independientes empleadas son de dos tipos: intrasujeto e intersujeto.

Las variables independientes intersujeto han sido:

- *Semántica-1*
- *Semántica-2, y*
- *Relación*

La variables *Semántica-1* y *Semántica-2* tienen cada una cuatro niveles que corresponden a las cuatro estructuras semánticas detectadas en problemas verbales simples de estructura aditiva que son:

- ES1: CA = Cambio
- ES2: CO = Combinación
- ES3: CP = Comparación
- ES4: IG = Igualación

La variable independiente *Relación* es una variable compuesta de dos más simples: tipo de relación en primera etapa y tipo de relación en segunda etapa, ambas con dos niveles posibles: la suma (relación de aumento) y la resta (relación de disminución). Tiene por tanto cuatro niveles que son:

- R1: (A,A)
- R2: (A,D)

R3: (D,A)

R4: (D,D)

Al cruzar estas tres variables independientes entre sí se obtiene un diseño factorial de dimensión 4x4x4, que da lugar a una matriz de 64 casillas, que representan cada una de las 64 tareas-problema que han debido resolver los alumnos en una situación de diseño intrasujeto.

TABLA 3
EJEMPLOS DE PROBLEMAS CORRESPONDIENTES A LAS CUATRO
COMBINACIONES DE AUMENTO Y DISMINUCIÓN

		Segunda posición	
		Aumento (A)	Disminución (D)
Primera posición	Aumento (A)	María tiene 12 sellos de Francia y 7 de España. Compra 16 sellos de Grecia ¿Cuántos sellos tiene en total?	José tiene 18 bolas rojas y 7 bolas negras. Después de jugar una partida pierde 11 bolas ¿Cuántas bolas le quedan?
	Disminución (D)	José tiene 25 bolas rojas y negras, 7 son negras. Le regalan 4 bolas rojas ¿Cuántas bolas rojas tiene José?	Juan tiene 35 cromos, 7 son de animales y el resto de aviones. Regala a un amigo 12 cromos de aviones ¿Cuántos cromos de aviones le quedan?

Por último, añadimos dos nuevas variables independientes intersujeto que completan el diseño anterior, las variables COLEGIO y CURSO, que confiere al trabajo una estructura de diseño factorial mixto, donde se analizan los seis colegios en los que se aplicaron las pruebas y los tres niveles: 4º, 5º y 6º de Educación Primaria.

Decir también que la investigación se ajusta a las coordenadas propias de un diseño cuasi-experimental, ya que ni los colegios ni los alumnos se han elegido aleatoriamente.

Variables controladas

En los 64 problemas aritméticos de dos etapas hemos controlado las siguientes variables intervinientes:

- el tipo de números: naturales;
- el tamaño de los números: inferiores a 60;
- el tamaño del resultado: inferior a 60;

- el tipo de magnitud: discreta;
- la naturaleza de los agentes: personas;
- el contexto del enunciado: familiar al niño;
- la longitud de las oraciones (expresión gramatical mínima: sujeto, predicado y objeto directo);
- la situación de la pregunta en el enunciado: al final;
- el formato del enunciado: tres frases separadas por signos de puntuación. En las dos primeras frases se aportan los datos, manteniendo el mismo orden en el que luego hay que operar con ellos; en la tercera se plantea la pregunta.
- Una variable importante en la categorización de los problemas verbales simples es la *posición de la incógnita o cantidad desconocida en el esquema* correspondiente al problema verbal de una etapa. Esta variable la hemos mantenido constante en nuestro estudio.

Dado que aplicar a los mismos alumnos los 64 tipos de problemas al mismo tiempo presenta dificultades evidentes de orden práctico, elaboramos pruebas diferentes mediante una clasificación de los 64 problemas en grupos de 8 problemas; esto permitió elaborar 8 pruebas distintas. En la Tabla 3 hemos recogido una muestra de problemas incluidos en este estudio.

Procedimiento de aplicación de las pruebas

Las ocho pruebas escritas de lápiz y papel confeccionadas (Prueba A, Prueba B, Prueba C, Prueba D, Prueba E, Prueba F, Prueba G y Prueba H) se aplicaron a los alumnos de 4º, 5º y 6º de Educación Primaria de los seis colegios que intervinieron en el estudio. Las pruebas se aplicaron siguiendo el orden alfabético de las mismas, es decir, primero se aplicó la prueba A, después la prueba B, continuando así hasta aplicar en último lugar la prueba H.

Con la finalidad de evitar efectos indeseables producidos por la reiteración de aplicación de pruebas a los mismos niños, se realizó una programación espaciada en un periodo de tiempo del curso escolar 1993-94.

Cada una de las pruebas se ha aplicado en días distintos, procurando dejar entre prueba y prueba al menos uno o dos días de distancia, y evitando el fin de semana en medio. Los niños resolvieron los problemas de forma individual en su mesa habitual de estudio y trabajaron en silencio. Cada prueba dispuso de una hora para su realización.

Las pruebas han sido aplicadas por miembros del Seminario CIEM, directamente vinculados a la investigación. Previo a la entrega de las hojas con los problemas la persona del equipo encargada de aplicarlas daba instrucciones a los escolares sobre la forma en que debían proceder.

Criterios de corrección

Las respuestas escritas dadas por los niños a los problemas han sido catalogadas como correctas o incorrectas. Cuando un niño no ha dado respuesta a un problema se

ha catalogado como incorrecta. En las respuestas dadas por los niños, y puesto que esta investigación se refiere al rendimiento en la fase de comprensión, la decisión de si un alumno ha comprendido o no un problema se ha basado en si el niño ha elegido o no operaciones que permitan obtener la solución. Hemos dado por correctas soluciones en las que el sujeto ha elegido una operación adecuada aunque se haya equivocado al operar.

Fiabilidad del conjunto total de ítems

Al ser los ítems dicotómicos (acierto-fracaso), hemos calculado la fiabilidad del conjunto de los 64 problemas mediante la fórmula KR-20 de Kuder Richardson que mide la consistencia interna del conjunto de ítems. El valor que hemos obtenido para la KR-20 ha sido de 0,96 que es un valor muy alto y garantiza la fiabilidad de los datos recogidos.

RESULTADOS

En este apartado presentamos el análisis cuantitativo de las respuestas de los niños en función del colegio, del curso y de la variable relación. La variable dependiente utilizada ha sido el número de respuestas correctas expresado en porcentaje de aciertos. Para analizar estadísticamente los resultados hemos aplicado un análisis de la varianza mixto $6 \times 3 \times 4$ (colegio \times curso \times relación) al porcentaje de respuestas correctas, con medidas repetidas en el último factor, en el que el colegio y el curso actúan como factores inter-sujetos y la variable de tarea «relación» actúa como factor intra-sujetos.

El ANOVA ha sido realizado con el paquete estadístico SPSS/PC+ y muestra (véase Tabla 4), en primer lugar, que son significativos los efectos principales del COLEGIO ($F = 14.677$, $p = 0.0$), CURSO ($F = 145.584$, $p = 0.0$) y RELACIÓN ($F = 18.377$, $p = 0.0$). En segundo lugar, que no hay interacción significativa COLEGIO \times CURSO, pero sí son significativas las interacciones COLEGIO \times CURSO ($F = 6.110$, $p = .000$) y CURSO \times RELACIÓN ($F = 3.230$, $p = .004$).

Las puntuaciones globales de los colegios, expresadas en porcentajes medios de resolución que oscilan entre 79.79 y 90.09. Realizadas «comparaciones múltiples» por el método de Scheffé ($p=0.05$), revelan que sólo hay diferencias significativas entre el colegio de menor rendimiento y todos los demás.

Con respecto al factor CURSO, para detectar entre qué parejas de cursos hay diferencias significativas hemos aplicado el método de Scheffé al nivel de significación del 5%. Los resultados obtenidos muestran que hay diferencias significativas entre todas las parejas de cursos posibles: 4º vs. 5º, 4º vs. 6º, y 5º vs. 6º. En la Ilustración 2 se observa que el rendimiento medio por curso en los problemas aritméticos verbales de dos etapas aumenta con la edad escolar. El incremento es progresivo, pero se da un salto mayor de 4º (76,5%) a 5º (88,6), que de 5º a 6º (93,8%). En estos cursos se produce, por tanto, un desarrollo de la capacidad de resolución de los problemas aditivos de dos etapas.

TABLA 4
RESULTADOS DEL ANÁLISIS DE LA VARIANZA EN UN DISEÑO FACTORIAL
(6 X 3 X 4) Y CON LA PROPORCIÓN DE RESPUESTAS CORRECTAS COMO
VARIABLE DEPENDIENTE (n=1036)

Fuente de Variación	Suma de Cuadrados	GL	Medias de Cuadrados	F	Signif. de F
Efectos Principales	144490.324	10	14449.032	42.123	0.0
COLEGIO	25172.985	5	5034.597	14.677	0.0
CURSO	99876.301	2	49938.151	45.584	0.0
RELACIÓN	18911.008	3	6303.669	118.377	0.0
Interacciones de 2 vías	33497.107	31	1080.552	3.150	.000
COLEGIO CURSO	20960.070	10	2096.007	6.110	.000
COLEGIO RELACIÓN	5821.088	15	388.073	1.131	.322
CURSO RELACIÓN	6648.300	6	1108.050	3.230	.004
Interacciones de 3 vías	7415.816	30	247.194	.721	.866
COLEGIO CURSO RELACIÓN	7415.816	30	247.194	.721	.866
Explicada	185403.247	71	2611.313	7.613	0.0
Residual	673688.876	1964	343.019		
Total	859092.123	2035	422.158		

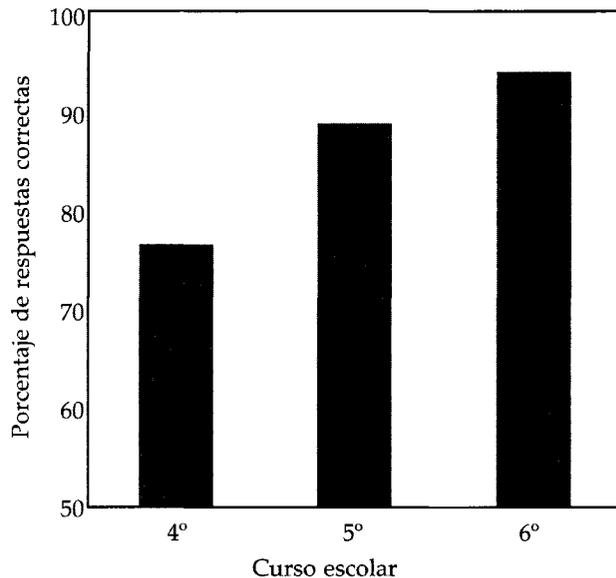


Ilustración 2.

La Ilustración 3 muestra los porcentajes de aciertos de los cuatro niveles de la variable RELACIÓN. Hay un escalonamiento suave en los índices de dificultad de las cuatro combinaciones que constituyen esta variable. Para detectar si hay diferencias significativas entre los índices de dificultad de los cuatro valores de la variable RELACIÓN hemos realizado «comparaciones múltiples» *a posteriori* por el método de Scheffé ($p = 0.05$). Con esta prueba hemos obtenido que la combinación de relaciones (D,D) difiere significativamente de la (A,A) y de la (A,D). También hemos obtenido que la combinación de relaciones (-,+), difiere significativamente de la combinación (A,A). Con estos resultados se obtienen tres parejas de combinaciones homogéneas cuyos rendimientos son muy similares entre sí, que son de mayor a menor dificultad:

- la clase 1 formada por las combinaciones (D,D) y (D,A)
- la clase 2 formada por (D,A) y (A,D), y
- la clase 3 formada por (A,D) y (A,A).

Como puede observarse hay diferencias muy pequeñas entre los porcentajes de la combinación de relaciones (D,A) y de la combinación (D,D). Debido a ello, y a que las diferencias entre la combinación (A,A) y la combinación (A,D) son también pequeñas, pensamos que, con respecto a la combinación de relaciones, la primera relación tiene un papel más decisivo que la segunda en el proceso de comprensión del problema. Esto es una conclusión que habría que analizarla estadísticamente con un diseño en el que la primera relación y la segunda se incluyeran como variables independientes.

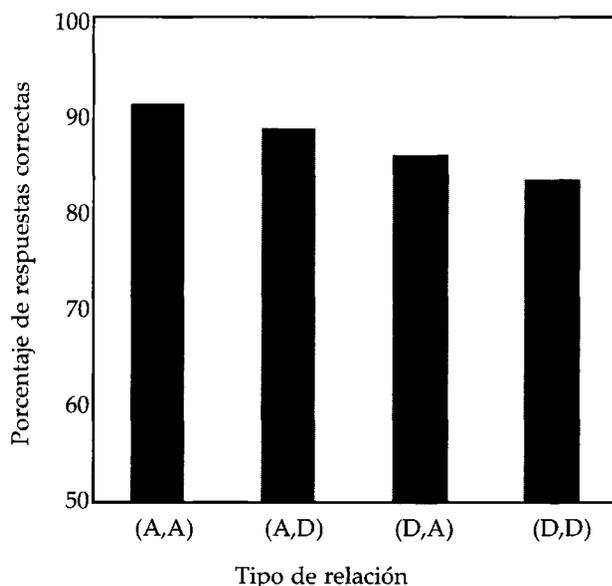


Ilustración 3. Porcentajes de aciertos de los cuatro niveles de la variable relación.

Interacción CURSO x RELACIÓN

Los resultados del análisis de la varianza muestran un efecto significativo de interacción entre el curso y la variable relación. El análisis de la interacción *curso x relación* lo hemos realizado mediante el estudio de los efectos de cada una de las variables según los distintos niveles de la otra y realizando en cada caso comparaciones múltiples. En primer lugar, hemos analizado el comportamiento de los tres cursos por separado y hemos obtenido que, en todos ellos, hay efecto significativo de la variable relación sobre el rendimiento de los niños. Se mantiene pues lo obtenido de manera general para la variable *relación*. Realizadas las comparaciones múltiples entre los niveles del factor relación por cada uno de los cursos, encontramos que en cuarto y en quinto las combinaciones (A,A) y (A,D) difieren significativamente de las (D,A) y (D,D), mientras que en sexto curso la combinación (A,A) difiere de las combinaciones (D,A) y (D,D). Así pues, en 4º y 5º curso hay dos clases homogéneas disjuntas: la primera, de dificultad menor, formada por las combinaciones que empiezan con aumento (A): (A,A) y (A,D), y la segunda, con una dificultad mayor que está formada por las combinaciones que empiezan por disminución (D): (D,A) y (D,D). En sexto curso, se mantiene la clase primera, pero a la segunda se incorpora también la (A,D) y por tanto las dos clases no son disjuntas.

En segundo lugar, hemos analizado el desarrollo que experimentan los niños de 4º, 5º y 6º de manera individual para cada uno de los niveles de la variable relación. Los resultados del análisis muestran que hay efecto significativo de la variable curso en cada uno de los cuatro niveles de la variable relación, al nivel del 5%. Las comparaciones múltiples efectuadas entre 4º, 5º y 6º curso han dado que hay diferencias significativas entre las tres parejas posibles 4º vs. 5º, 4º vs. 6º y 5º vs. 6º. Estos resultados son exactamente los obtenidos para la variable curso en el análisis global.

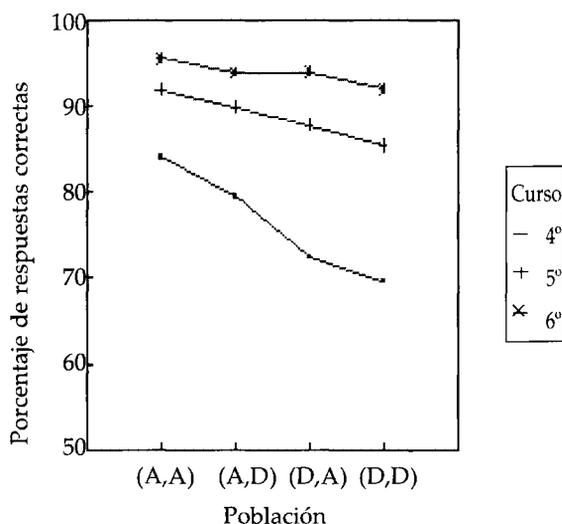


Ilustración 4. Porcentajes de aciertos asociados al tipo de relación según el curso.

Resumiendo, podemos decir que la interacción entre las variables RELACIÓN y CURSO no altera el orden de dificultad encontrado entre los niveles de cada variable por separado. Podemos observar en la Ilustración 4 que el avance por curso se produce en los cuatro niveles de la variable relación. Así mismo, el orden de dificultad de los niveles de la variable relación es el mismo en los tres cursos, salvo que se produce un mayor acercamiento entre los índices de dificultad de las cuatro combinaciones al avanzar en el curso escolar. Los contrastes estadísticos han dado dos grupos de problemas homogéneos, pero aunque no sea a nivel significativo hay que resaltar que la combinación (D,D) tiene el mayor índice de dificultad, seguida de (D,A), a continuación viene (A,D) y la más fácil (A,A) y esto ocurre en 4º, 5º y 6º.

CONCLUSIONES

El rendimiento medio obtenido en porcentaje por la muestra empleada de 509 alumnos de 4º, 5º y 6º ha sido del 86,98%. Esto, en principio, lleva a considerar que este tipo de problemas no plantean una dificultad alta de comprensión para este colectivo de alumnos, tomados globalmente. A la vista de este dato puede pensarse que todos los alumnos de la muestra están en el mismo nivel de comprensión en todas las tareas que se han propuesto. El análisis estadístico muestra que no es así, aunque unos aspectos son más influyentes que otros. Veámoslo con detalle. Nuestra intención al incorporar la variable *colegio* en el estudio ha sido controlar en lo posible la fenomenología inherente a distintos colegios, que pudiera incidir en los otros aspectos que estamos estudiando. El análisis estadístico ha puesto de manifiesto que de los seis colegios que participaron en la experiencia cinco de ellos tienen un rendimiento homogéneo. El otro colegio tiene un rendimiento significativamente menor que estos cinco. A la vista de estos resultados sorprende el hecho de que, aún siendo colegios muy diferentes estructuralmente en lo que se refiere a variables de contexto tales como la ubicación geográfica, el nivel social o incluso la diversidad de metodologías y tratamientos curriculares seguidos por los profesores, sin embargo, los rendimientos globales son más homogéneos de lo que, a priori, podría esperarse.

Por razones de ser los grupos de la muestra grupos naturales es también lógico que haya una cierta variabilidad en el nivel de rendimiento de los cursos 4º, 5º y 6º según los distintos colegios. Por ello, la interacción *colegio x curso* es hasta cierto punto esperada y deseada por nosotros, ya que una muestra heterogénea puede darle mayor generalidad a los resultados que hemos obtenido en los otros aspectos en que estamos interesados.

Un hallazgo importante en esta investigación sobre problemas aritméticos verbales de dos etapas ha sido que la dificultad de los niños para seleccionar las operaciones adecuadas para resolver un problema está afectada por el año escolar en el que se encuentra. Con respecto al año escolar, se produce un desarrollo significativo en los niños a lo largo de estos tres cursos en la capacidad de resolución de problemas aritméticos verbales aditivos de dos etapas. La explicación a este hecho puede ser doble: Por un lado puede deberse a la instrucción recibida y por otro al desarrollo psicológico que se produce en estas edades. Puesto que no ha habido en estos colegios

un objetivo o intención explícita por mejorar la capacidad de resolución de los problemas aritméticos de dos etapas a lo largo de 4°, 5° y 6° de Primaria, ni se ha llevado una metodología tendente a mejorar esa capacidad, pensamos que la explicación a esta maduración que experimentan los niños es consecuencia indirecta de la instrucción escolar recibida y/o al desarrollo cognitivo.

Es lógico, pensar que el niño con la edad tiene una mayor capacidad de resolución de problemas, pero esto por sí mismo no permite decidir en un momento determinado si es capaz de resolver un tipo o una clase concreta de problemas. Encontrar qué problemas son adecuados a un nivel escolar, cuáles les superan su capacidad de resolución y cuáles han sido superados, es un conocimiento útil para el desarrollo del currículo de matemáticas. Al finalizar la educación primaria hemos encontrado que un alto porcentaje de los niños de 6° resuelven correctamente esta clase de problemas. Es, pues, una capacidad que se ha adquirido durante la educación primaria y lo interesante es analizar cuáles fueron las etapas claves en este aprendizaje. Si en 6° curso pensamos que el alumno en general ha desarrollado la capacidad, en 4° todavía hay aproximadamente un 25% de los alumnos que tienen dificultades para comprender los problemas de dos etapas. De 4° a 5° curso se produce un incremento considerable en su capacidad de resolución de los problemas aritméticos verbales de dos etapas de estructura aditiva, incremento que también se produce en el paso de 5° a 6°, pero en menor medida. Estos datos nos hacen conjeturar que el salto de 3° a 4° debe ser aún mayor que los anteriores, y por tanto, consideramos que los cursos 3° y 4° son claves en el desarrollo de la capacidad de resolución de los problemas verbales aditivos de dos etapas.

También hay que reseñar que esta evolución durante los años de escolaridad se ve afectada por el colegio, en el sentido de que en unos colegios se produce antes que en otros. Pero lo que hay que dejar claro es que en todos se produce y que en sexto no hay diferencias significativas de unos colegios a otros. Todo ello referido a los colegios de la muestra.

Las combinaciones de relaciones de aumento o disminución no afectan a este desarrollo. Dicho de otra manera, el desarrollo que experimentan los niños a lo largo de los cursos, se produce en las cuatro combinaciones. Lo que varía de una combinación a otra es la intensidad del desarrollo experimentado, produciéndose un mayor avance en las combinaciones menos desarrolladas en los niños hasta el momento: (D,A) y (D,D), y un menor avance en aquellas que están más consolidadas de cursos anteriores: (A,A) y (A,D).

El segundo hallazgo importante que hemos encontrado es que las distintas combinaciones de las relaciones de aumento o disminución afectan a la dificultad del problema. Los problemas de dos etapas constan de dos relaciones que pueden ser de aumento o de disminución respectivamente. Las cuatro clases de problemas que surgen según que la relación sea de aumento o de disminución, colocados en orden de dificultad creciente quedan así:

(A,A), (A,D), (D,A) y (D,D)

siendo este último, el (D,D) el de mayor dificultad.

No es fácil comparar este resultado con otros que pueden parecer similares y que hayan sido obtenidos en investigaciones previas. Y ello por una razón sencilla: El planteamiento teórico ha cambiado. Las primeras investigaciones de Berglaund-Gray (1938, 1939), Berglaund-Gray & Young (1940), y Young & McIsaac (1941) investigan el efecto de invertir la secuencia de operaciones. En estas investigaciones primero está la operación, y el problema se supone que es de sumar, restar multiplicar o dividir. En la situación actual las operaciones han perdido su papel como criterios clasificatorios *a priori* de los problemas aritméticos verbales y han sido relegadas a su verdadera función *a posteriori* que es el de herramientas con las que se pueden resolver los problemas verbales aritméticos. Esto tiene implicaciones para decidir si un problema es de uno u otro tipo. Pese a estos distanciamientos teóricos los resultados que hemos obtenido son coincidentes para los casos comunes a ambas investigaciones. Los autores anteriores concluyen que la substracción fue más difícil como primer proceso que la adición. Esto mismo hemos obtenido nosotros, pero además lo hemos obtenido en la totalidad de la muestra y en cada uno de los cursos. Lo anterior nos hace concluir que, en un problema de estructura aditiva compuesto de dos relaciones, la relación de disminución en el primer paso es más difícil que las relaciones de aumento en primer paso.

Recordamos que para obtener la variable relación hemos combinado dos relaciones simples de aumento o disminución que ocupan en el enunciado del problema la primera y la segunda posición, respectivamente. Según los resultados obtenidos la influencia de la relación de aumento o disminución en segunda posición, tiene un peso menor sobre la dificultad del problema que la que tiene en la primera posición. De hecho, no hemos obtenido unas diferencias significativas según la relación en la segunda posición. Lo que sí ocurre es que la relación de disminución en la segunda posición refuerza la dificultad que provoca la relación de disminución en la primera posición. Esto ocurre en la muestra total y en cada uno de los cursos por separado, con la única salvedad de que los índices de dificultad se mueven en un intervalo menor.

REFERENCIAS

- BARNETT, J. (1980). The study of syntax variables. En G.A. Goldin y C.E. McClintock (Eds.), *Task Variables in Mathematical Problem Solving*. Philadelphia, Pennsylvania: The Franklin Institute Press.
- BELL, A.; GREER, B.; GRIMISON, L. y MANGAN, C. (1989). Children's performance on multiplicative word problems: Elements of a descriptive theory. *Journal for Research in Mathematics Education*, 20, 434-449.
- BERGLAUND-GRAY, G. (1938). *The effect of process sequence on the interpretation of two-step problems in arithmetic*. Unpublished doctoral dissertation. University of Pittsburgh.
- BERGLAUND-GRAY, G. (1939). Difficulty of the arithmetic process. *Elementary School Journal*, November, 40, 198-203.
- BERGLAUND-GRAY, G. & YOUNG, R.V. (1940). The effect of process sequence on

- the interpretation of two-step problems in arithmetic. *Journal of Educational Research*, September, 34(1), 21-29.
- CARPENTER, T.P. y MOSER, J.M. (1982). The Development of Addition and Subtraction problem-solving skills. En T. P. Carpenter, J. M. Moser y T. A. Romberg (Eds.), *Addition and Subtraction: A cognitive perspective*. Hillsdale, N. J.: LEA.
- CARPENTER, T.P. y MOSER, J.M. (1983). The Acquisition of Addition and Subtraction concepts. En R. Lesh y M. Landau (Eds.), *Adquisition of Mathematics Concepts and Processes*. Orlando, Florida: Academic Press.
- CASTRO, E.; RICO, L. y GIL, F. (1992). Enfoques de Investigación en Problemas Verbales Aritméticos Aditivos. *Enseñanza de las Ciencias*, 10 (3), 243-253.
- CASTRO, E. (1995). *Niveles de comprensión en problemas verbales de comparación multiplicativa*. Granada: Comares.
- DURELL, F. (1928). Solving problems in arithmetic. *School Science and Mathematics*, 28(9), 925-935.
- FUSON, C. (1992). Research on Whole Number Addition and Subtraction. En D.A. Grouws (Ed.), *Handbook of Research on Mathematics Teaching and Learning*. New York: MacMillan Publishing Comp.
- GOLDIN, G. A. & McCLINTOCK, C.E. (Eds.) (1980). *Task Variables in Mathematical Problem Solving*. Philadelphia, Pensilvania: The Franklin Institute Press.
- GUTIÉRREZ, J.; MORCILLO, N.; RICO, L.; CASTRO, E.; CASTRO, E.; FERNÁNDEZ, F.; GONZÁLEZ, E.; PÉREZ, A.; SEGOVIA, I.; TORTOSA, A. y VALENZUELA, J. (1993). Problemas aditivos de dos etapas con igual operación y estructura semántica duplicada. Estudio preliminar en 5° de Primaria. *Actas VI JAEM*, Badajoz.
- HELLER, J.I. y GREENO, J.G. (1979). Information Processing analysis of mathematical problem solving. En R. Lesh, D. Mierkiewicz y M. Kantowski (Eds.), *Applied Mathematical Problem Solving*. Columbus, Ohio: ERIC/SMEAC.
- JERMAN, M. y REES, R. (1972). Predicting the relative difficulty of verbal arithmetic problems. *Educational Studies in Mathematics*, 4, 306-323.
- MORCILLO, N.; CASTRO, E.; RICO, L.; CASTRO, E.; FERNÁNDEZ, F.; GONZÁLEZ, E.; GUTIÉRREZ, J.; PÉREZ, A.; SEGOVIA, I.; SERRANO, M.; TORTOSA, A. y VALENZUELA, J. (1993). Dificultad debida al orden de operaciones en Problemas Aditivos de Dos Etapas con estructura semántica duplicada. Estudio preliminar en 5° de Primaria. *Actas VII Jornadas Andaluzas de Educación Matemática «Thales»*. Sevilla.
- NESHER, P. (1982). Levels of description in the analysis of addition and subtraction word problems. En T. P. Carpenter, J. M. Moser y T. A. Romberg (Eds.), *Addition and Subtraction: A Cognitive Perspective*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- NESHER, P. (1991). Two-Steps Problems, Research Finding. En F. Furinghetti (Ed.), *Proceedings Fifteenth PME Conference*, Vol. III, pp. 65-71. Assisi, Italia.
- NESHER, P. y HERSHKOVITZ, S. (1994). The role of schemes in two-step problems: analysis and research findings. *Educational Studies in Mathematics*, 26, 1-23.
- PUIG, L. y CERDÁN, F. (1988). *Problemas aritméticos escolares*. Madrid: Editorial Síntesis.
- RICO, L. y OTROS (1985). *Investigación «Granada-Mats»*. Granada: Instituto de Ciencias de la Educación, Universidad de Granada.

- RICO, L. y OTROS (1988). *Didáctica activa para la resolución de problemas*. Granada: Departamento Didáctica de la Matemática de la Universidad de Granada.
- RICO, L. y OTROS (en prensa). Categorías de problemas aditivos de dos etapas. *Educación Matemática*.
- SHALLIN, V.L. y BEE, N.V. (1985). *Structural differences between two-step word problems*, presentado en el Meeting de la American Educational Research Association.
- SHERARD, W.H. (1974). *The effect of arithmetical operations on the difficulty levels of verbal problems* (Doctoral dissertation, George Peabody College for Teachers, 1974). Dissertation Abstracts International, 1974, 35, 2895B. (University Microfilm No. 74-29, 189).
- SUPPES, P.; LOFTUS, E. y JERMAN, M. (1969). Problem-Solving on a computer-based teletype. *Educational Studies in Mathematics*, 2, 1-15.
- YOUNG, R. & MCISAAC, J. (1941). The sequence of processes affects the pupil's interpretation of verbal problems in arithmetic. *Education*, 61, 488-491.

PRUEBAS ESTANDARIZADAS Y EVALUACIÓN DEL RENDIMIENTO: USOS Y CARACTERÍSTICAS MÉTRICAS

J.M. Jornet Meliá y J.M. Suárez Rodríguez¹

RESUMEN

En este artículo presentamos algunas reflexiones sobre el uso de las pruebas estandarizadas para la evaluación del rendimiento. Se propone una tipología de pruebas cuyos componentes son: pruebas como Indicadores de resultados, pruebas de certificación y de admisión, pruebas de dominio, pruebas de clase y pruebas individualizadas. Para cada tipo de prueba se revisan las propiedades métricas que se derivan de los objetivos, características y finalidad de las pruebas.

ABSTRACT

In this article we present some reflections about the use of the standardized tests for achievement evaluation. It is proposed a test typology whose components are: tests as outcome indicators, certification and admission tests, domain tests, classroom tests and tailored tests. The measurement properties derived from the tests objectives, characteristics and purpose are reviewed for each kind of test.

¹ Dpto. Mètodes d'Investigació i Diagnòstic en Educació. Universitat de València (Estudi General). Avda. Blasco Ibáñez, 21. 46010-Valencia. Tél. y Fax: 96/3864430. E-mail: Jesus.M.Jornet@uv.es / Jesus.M.Rodriguez@uv.es

INTRODUCCIÓN

La utilización de pruebas Estandarizadas en el ámbito Educativo es muy frecuente. En la literatura se utilizan cotidianamente términos que denominan diversos tipos de instrumentos que aluden a este tipo de pruebas: Tests Referidos al Dominio, Tests Referidos a Objetivos, Tests de Competencia, Tests de Certificación, Tests de Dominio, Tests referidos al Criterio, etc.... No obstante, en nuestro medio sociocultural, su uso es más bien escaso y, ciertamente, son pocas las pruebas estandarizadas de Rendimiento que se hayan desarrollado en nuestro país y para nuestro sistema educativo. Las razones que están a la base de este fenómeno pueden ser diversas, pero desde nuestro punto de vista, los usos equívocos de este tipo de pruebas han arraigado la concepción de que son poco útiles a efectos evaluativos y que, en todo caso, su uso está indefectiblemente ligado a corrientes pedagógicas que atienden poco a las características de los individuos. Obviamente, esta posición que atribuimos a buena parte de los detractores de las Pruebas Estandarizadas está simplificada y, probablemente, sería matizada de diversas formas, sin embargo quizá es la posición más generalizada entre ellos.

Desde nuestro punto de vista, el problema normalmente radica en que se pretende de las pruebas estandarizadas usos e interpretaciones para los que normalmente no han sido construidas y, en ocasiones, se desarrollan con esquemas de elaboración que han sido diseñados metodológicamente para objetivos evaluativos diferentes.

Generalmente, la inadecuación de las pruebas está en parte debida a que los criterios de construcción de pruebas se presentan de manera indiferenciada. Aunque son pocos los Modelos de Medida disponibles, las variaciones en su aplicación pueden ser múltiples. En la adaptación precisa de estos elementos radica buena parte de la calidad de las pruebas estandarizadas. Por adaptación nos referimos en este caso al ajuste de los métodos, procedimientos y técnicas de elaboración a las características concretas de la prueba que se desea construir. Estas características son, a su vez, consecuencia del compromiso de diversos factores como son: el objeto de medida, finalidad / uso de la prueba y las características de las personas a las que se desea evaluar a través de la prueba. En este contexto, puede ser de interés realizar algunas reflexiones acerca de los componentes generales de actuación en la elaboración de pruebas estandarizadas, que permitan un mayor aprovechamiento de éstas para los procesos evaluativos.

DIMENSIONES DE CLASIFICACIÓN DE LAS PRUEBAS ESTANDARIZADAS

En la literatura especializada en Medición y Evaluación se identifican una gran cantidad de términos referidos a pruebas estandarizadas. Ante esta diversidad es conveniente determinar algunas dimensiones que nos permitan abordar su clasificación. Entre estas dimensiones nos centraremos en aquéllas que están relacionadas con los componentes métricos o metodológicos de su elaboración. En este caso, el grado de estandarización no entra a formar parte de las dimensiones de clasificación, dado que es una característica constante en todas las pruebas a las que aquí nos referiremos.

Como señalamos en la introducción a este artículo, los tipos de pruebas devienen

de la confluencia de diversos factores. En ellos, podemos identificar dimensiones de definición que afectan a la construcción de las pruebas. Generalmente estas dimensiones son de carácter bipolar y definen un eje en el cuál pueden situarse las características de cada prueba de forma gradual. Así, comentaremos brevemente cada dimensión identificándola con sus polaridades. En todos los casos, al entenderse que estas dimensiones son graduales, cabe establecer un elemento de clasificación intermedio, que corresponden a “enfoques mixtos”, que por no ser reiterativos obviaremos en la exposición subsiguiente de dimensiones.

- *Características del Objeto de Medida.* Las características del *Dominio Educativo*² a que va dirigido la prueba es un elemento de definición básico que condiciona, desde los elementos de Validez, el conjunto del desarrollo de la prueba. El Dominio Educativo constituye el *Universo de Medida* desde el que se extraen los componentes de la prueba y al que se pretende representar desde ella. Sin entrar en los factores específicos de los tipos de contenidos educativos, las dimensiones a tener en cuenta para orientar el desarrollo de las pruebas son las siguientes:
 1. Amplitud del Dominio Educativo (Dominios amplios vs. reducidos).
 2. Límites del Dominio Educativo (Dominios con límites difusos/no-finitos vs. Dominios con límites concretos/finitos).
 3. Dimensionalidad del Dominio Educativo (Dominios Multidimensionales vs. Unidimensionales).
- *Características de la población a que va dirigida la prueba.* Afecta fundamentalmente a la elección del Modelo de Medida y la selección de indicadores que permitan el análisis adecuado del ajuste de las características de la prueba a las de la población. Las dimensiones más relevantes son:
 4. Amplitud de la población (Población extensa vs. Reducida).
 5. Grado de diversidad de la población (Población Heterogénea vs. Homogénea).
- *Finalidad y uso de la prueba.* La Validez no es en sí misma una característica imputable a una prueba, es más bien el uso que se pretende realizar de las puntuaciones de ella derivadas lo que debe analizarse como elemento de validación (Hambleton; 1984). Así, la utilización que se desee realizar de la prueba tiene consecuencias desde la definición del Dominio Educativo hasta el establecimiento de Estándares de puntuación. Las dimensiones más importantes que pueden identificarse en este punto son:
 6. Decisiones asociadas al uso de la prueba (De carácter Formativo vs. Sumativo).
 7. Unidades sobre las que se pretenden tomar decisiones (Individuos vs. Grupos).
- *Características del tipo de Interpretación de puntuaciones.* Como en el caso anterior, estos elementos afectan a todo el desarrollo de la prueba. La dimensión central a que pueden reducirse estas características es:
 8. Tipo de Estándar de referencia (Normativo vs. Criterial).

2 Por Dominio Educativo nos referimos al conjunto de objetivos, contenidos, actividades y tareas que constituyen el objeto de la educación, sea en general sea en un programa concreto (Jornet y Suárez, 1989a).

El cruce de estas dimensiones puede servir para identificar las características de diversos tipos de pruebas estandarizadas y orientar los componentes específicos de sus procesos de elaboración. Una propuesta, aunque no exhaustiva, de tipología de pruebas estandarizadas se recoge en el Cuadro 1. Los elementos metodológicos y los aspectos que entendemos más relevantes en su construcción los comentaremos a continuación.

PRUEBAS DE AMPLIO ESPECTRO

En este apartado revisamos las características y usos de pruebas de Rendimiento que se orientan a la evaluación de grandes Áreas o Dominios Educativos. En esta categoría incluimos las pruebas que se utilizan como Indicadores de Resultados para

CUADRO 1
TIPOS FUNDAMENTALES DE PRUEBAS ESTANDARIZADAS VALORADOS
RESPECTO A OCHO DIMENSIONES BÁSICAS DE CARACTERIZACIÓN

Tipo de Prueba Estandarizada	DIMENSIONES DE VALORACIÓN							
	Amplitud del Dominio Educativo	Límites del Dominio Educativo	Dimen-sionalidad ³	Amplitud de la Población	Grado de diversidad de la Población	Decisiones Asociadas	Unidades sobre las que se decide	Tipo de Estándar
DE AMPLIO ESPECTRO: • Indicadores De Resultados • Certificación • Admisión	Amplio	No-finitos Difusos	Multidimen-sionales	Amplia / Muy amplia	Heterogénea	Formati-vas/ Sumativas	Grupos Individuos	Mixto: Normativo Normativo Criterial
DE NIVEL O DOMINIO	Mixto	Mixto	Multidimen-sionales	Intermedia / Amplia	Mixto	Sumativas	Individuos	Criterial
DE CLASE	Reducido	Finitos Concretos	Unidimen-sionales	Reducida / Muy reducida	Homogé-neo	Formati-vas/ Sumativas	Individuos	Criterial
DE PROPÓSITO DIAGNÓSTICO	Mixto	Finitos Concretos	Mixtos	Amplia/ Reducida	Heterogé-neo	Formati-vas	Individuos	Normativo Criterial
INDIVIDUALIZADAS	Reducido	Finitos Concretos	Unidimen-sionales	Reducida	Homogé-neo	Formati-vas/ Sumativas	Individuos y/o Grupos	Criterial

3 Hace referencia a las características originales del Dominio Educativo. Todas las pruebas es preciso adecuarlas a Universos Unidimensionales, por lo que en el caso de universos multidimensionales, se focalizan las pruebas sobre regiones específicas del Dominio.

el Análisis y/o Evaluaciones de Sistemas Educativos, Centros y Programas, pruebas de Certificación y pruebas de Admisión.

- ***Pruebas Estandarizadas como Indicadores de Resultados.***

La actividad evaluativa forma parte de la cultura de gestión de los Estados democráticos. En el ámbito pedagógico pueden observarse diversos modelos y enfoques de Evaluación de los sistemas educativos en los que confluyen indicadores de diferente índole.

Para la construcción de indicadores de resultados, parece claro que cuando se aborda el análisis de un Sistema Educativo, de un Programa o de un Centro, uno de los indicadores a tener en cuenta —aunque no de forma exclusiva— son los resultados esenciales del programa (De Miguel, et al., 1994; Pérez Juste y Martínez Aragón, 1989; Tejedor et al., 1994).

En estos contextos es necesario utilizar pruebas estandarizadas que “traduzcan” los niveles de competencia que en las diferentes disciplinas y materias, una sociedad asume como objetivo educativo.

En este sentido, no es posible abordar un análisis adecuado de un sistema o un Programa si no se cuenta con pruebas estandarizadas de probada fiabilidad y validez.

Así, buena parte de los modelos de evaluación de Sistemas Educativos basados en indicadores⁴ incorporan indicadores de resultados del aprendizaje de los alumnos sustentados sobre pruebas estandarizadas, diferenciándolos de las calificaciones escolares o de otros indicadores de síntesis (como las tasas de egresados) que suelen identificarse como Resultados del Sistema. En los campos de la evaluación de centros y de programas también resulta habitual la utilización de estas pruebas como indicadores. Mayor tradición, si cabe, tiene la utilización de pruebas de este tipo como indicadores para actuaciones evaluativas a la medida en muy diversos niveles educativos, ámbitos de referencia y objetivos (a partir de los servicios de instituciones como el ETS en USA, el APU para Inglaterra, Gales e Irlanda del Norte; o el CITO en el contexto holandés-alemán).

¿Qué componentes están implicados en la elaboración de estas pruebas?

La definición del Dominio a que se refieren estas pruebas debe realizarse por un Comité de Expertos en la Materia objeto de evaluación, apoyados por especialistas en Medición y Evaluación como asesores metodológicos. Los problemas que deben enfrentar este tipo de Comités son variados y de su adecuada solución depende en

4 Existen sistemas de indicadores que permiten la comparabilidad entre diversos países y utilizan un número relativamente reducido de pruebas como es el caso de la OCDE (CERI/INES; 1995) o la Comunidad Europea (West et al., 1995). A un nivel intermedio se encuentran los programas desarrollados por la IEA (Postlethwaite, 1987), finalmente existen otros más completos en cuanto a la información que emplean sobre productos educativos como el sistema federal USA (SSPEI, 1991) o algunos otros sistemas más recientes que se están impulsando en el ámbito iberoamericano (por ejemplo la propuesta de Martínez Rizo; 1996). En nuestro país, el Instituto Nacional de Calidad y Evaluación (INCE) está desarrollando pruebas como indicadores de resultados del sistema educativo, habiéndose comenzado a publicar los primeros resultados (Gil, González y Suárez, 1995; INCE, 1996).

buena medida la validez y utilidad de las pruebas resultantes. Pasamos a revisar brevemente los elementos que caracterizan métricamente el desarrollo de estas pruebas y, en consecuencia, condicionan su uso.

El *Dominio Educativo*, como Universo de Medida, suele ser muy amplio, referido generalmente a una materia o disciplina considerada en función de los objetivos terminales de todo un período educativo (por ejemplo, las Matemáticas o el Lenguaje en la Primaria, o al final de la Secundaria) o, a lo sumo, se focalizan sobre grandes dimensiones de las mismas (Medida o Algebra, Comprensión Lectora...). Además, los límites del Dominio suelen ser difusos, dado que junto a la amplitud del mismo, se suma como dificultad añadida en la definición del Dominio el hecho de que suelen ser pruebas cuya finalidad es la evaluación en una gran población de sujetos, afectados por muy diversas aplicaciones de los Diseños Curriculares de referencia, desarrollados sobre diferentes modelos didácticos.

Este hecho conlleva que nos encontremos ante Dominios en la práctica no-finitos, en los que no es posible establecer una estrategia de muestreo probabilístico desde el Universo de Medida para configurar la Prueba⁵. De esta forma, la definición del Dominio debe realizarse sobre productos educativos esenciales, equiprobables a través de diferentes subpoblaciones y que mantengan sus parámetros fundamentales invariantes en las subpoblaciones identificables en la población.

Una dificultad adicional en la definición de este tipo de Dominios reside en que, por la amplitud del Universo de Medida, generalmente no son unidimensionales y están apoyados en constructos teóricos débiles⁶, con pocas evidencias de validación. En este tipo de pruebas, los avances más claros corresponden a estructuraciones dimensionales por el nivel cognitivo que implican las tareas-ítems.

Así, si bien la base de desarrollo de la definición del Dominio recae sobre el juicio de expertos, también es cierto que su comprobación se sustenta fundamentalmente sobre comprobaciones estructurales basadas en el análisis empírico de los resultados de las pruebas. En este sentido, un problema adicional que encontramos en estos desarrollos es que la comprobación de la Unidimensionalidad⁷ es difícil, pues, aunque existan propuestas metodológicas específicas para este tipo de análisis con variables dicotómicas, ciertamente los resultados son insatisfactorios dado que las dimensiones en muchas ocasiones se identifican por la dificultad de los ítems y no son interpretables desde los contenidos.

Por ello, en las estrategias de desarrollo de este tipo de pruebas es especialmente

5 Ante la imposibilidad de concretar todos los miembros del Universo de Medida, se pasa a utilizar estrategias de representación fundamentadas en tipologías básicas o elementos clave en la estructura del Dominio.

6 Es decir, no suelen estar desarrollados a partir de una teoría de aprendizaje que, de manera holística, globalice y de sentido a la estructuración y funcionalidad de la prueba.

7 Siendo este un supuesto básico sobre el que se sustenta la medida desde cualquiera de los modelos actualmente existentes y que se enraiza en la información que se extrae en cada elemento de la prueba (Hambleton y Swaminathan, 1985; Osterlind, 1992). Aunque se han planteado algunas alternativas para superar este problema, como los trabajos de Reckase (1979) o Samejima (1974), hasta el momento no se pueden considerar como auténticas opciones disponibles.

importante el trabajo realizado por el Comité de expertos, anticipando la estructura teórica posible del Dominio e identificando regiones en el mismo, de forma que para cada uno de los subdominios se configuren pruebas específicas.

Junto a los problemas teóricos contemplados en líneas anteriores, es preciso considerar en la definición del Dominio de estas pruebas, algunos aspectos de orden práctico ligados a la funcionalidad de las mismas. Si se desea desarrollar pruebas que sirvan como Indicador de Resultados y utilizarlas en evaluaciones sucesivas, no es útil configurar una sola prueba⁸. La alternativa de elaborar Formas Paralelas es compleja y difícil de lograr. En este marco, la creación de Bancos de Reglas de Generación de Elementos⁹ —lo más deseable— o Bancos de Ítems —lo más frecuente— se configuran como alternativas que nos acercan a la posibilidad de disponer de pruebas aleatoriamente paralelas o al menos equivalentes.

En este contexto de desarrollo es especialmente importante el trabajo que realiza el Comité de expertos en cuanto a la formulación y revisión de Ítems. Así, un aspecto crucial en la elaboración de este tipo de pruebas es el Análisis Lógico de Ítems que se centra sobre diferentes elementos.

En primer lugar, respecto a la selección inicial de ítems, es conveniente basarla en dos dimensiones del contenido: a) la importancia de los ítems, y b) su dificultad teórica. Esta consideración de dos dimensiones facilita que los ítems sean propuestos y seleccionados desde la idea, antes señalada, de que representen conocimientos o habilidades esenciales, cubriendo a su vez diferentes estratos de dificultad. Desde esta estrategia se evita que la selección de ítems se contamine con la idea de “mínima competencia”, la cual, en ocasiones, es entendida como expresión de la dificultad —y no de la importancia—: este hecho constituye una desviación frecuente del trabajo de estos comités.

Otro elemento a tener en cuenta en la formulación de los ítems es su calidad técnica, la cual es necesario revisar inicialmente por procedimientos lógicos y, posteriormente, basándose en resultados de ensayos piloto. En el caso de utilización de ítems de Elección Múltiple, el análisis de distractores debe constituir un trabajo central de este aspecto. Así, es tanto más importante el control y la anticipación de la dificultad y la adivinación desde la formulación de los ítems, que desde el análisis empírico de resultados.

Junto a estos elementos, el análisis del Sesgo supone un aspecto clave para la validez de este tipo de pruebas. Debe tenerse en cuenta que éstas son pruebas destinadas a la Medición y Evaluación de un Dominio en una amplia población de referencia. De esta manera, es conveniente anticipar en el análisis lógico las variables que podrían

8 Una sola aplicación de una prueba de estas características puede inhabilitar su uso. Una vez es conocida una prueba de este tipo puede ser utilizada como objeto directo de aprendizaje.

9 Son procedimientos que concretan de forma unívoca al ítem de forma que su escritura se torna automática, entre ellos los más destacados son los que se recogen en Roid y Haladyna (1982). Aunque lentamente, los procedimientos han ido evolucionando para tratar de hacer frente a la evaluación de la actuación en tareas complejas, incrementando el nivel cognitivo de la evaluación (por ejemplo, los conjuntos de ítems —ítem sets— desarrollados por Haladyna —1992—, o el modelado de ítems desarrollado por La Duca —La Duca et al. 1986—).

actuar como fuente de sesgo. No vamos a extendernos aquí en estos aspectos, dado que han sido anteriormente expuestos en esta Revista (Jornet y Suárez, 1990; Ordeñana, 1991); sin embargo, en este tipo de pruebas es básica la independización del sesgo, el cual puede provenir de diferentes fuentes como el Sexo, el ámbito territorial, los niveles socioculturales o socioeconómicos, la Lengua, etc...

El control de todos estos elementos se basa en dinámicas de juicio bien establecidas, en las que en la síntesis de resultados se utilizan procedimientos de análisis de la consistencia inter-jueces.

En este sentido, hay que destacar la necesidad de los procedimientos de juicio, el estudio de las dinámicas más adecuadas a seguir por parte de los Comités y de los tipos de análisis a realizar, tanto como elementos de síntesis de la opinión de los Comités, como para detectar jueces que aportan valoraciones extremas, etc.

La Validez de las pruebas se asienta sobre procesos de análisis teórico de los componentes de medida y la revisión lógica de todas las unidades¹⁰. Es fundamental reconocer el valor de los procesos de juicio en este ámbito, los cuales deberán ser atendidos adecuadamente. No obstante, un problema habitual que se contempla en este tipo de pruebas es que muchas de ellas se sustentan más sobre el Modelo de Medida utilizado que sobre el análisis teórico del Dominio Educativo. Por mencionar tan sólo uno de los elementos clave en que se asienta la Validez de Constructo. De hecho, el problema estriba en que los principios que se refieren a la validez y que están recogidos en cualquiera de los modelos de medida no son sino una parte relativamente reducida de los indicios que definimos como facetas de la Validez de Constructo (Angoff, 1988). Por ello las aportaciones de los modelos de medida deben entenderse en un plano instrumental dentro de la estrategia global de validación y no a la inversa.

¿Qué Modelo de Medida es más adecuado en este contexto?

Si se pretende utilizar estas pruebas sobre una población amplia, la base métrica necesariamente se encuentra en la Teoría de Respuesta al ítem que favorece el desarrollo de pruebas sobre parámetros invariantes de los ítems y que permiten una graduación adecuada de los ítems asociados con la habilidad general que mide la prueba (Hambleton y Swaminathan, 1985; Weiss y Yoes, 1991). No obstante, estos modelos han demostrado su adecuación, hasta el momento, con dos condiciones bastante precisas no existiendo un acuerdo generalizado sobre su utilización cuando alguna de ellas no se cumple. La primera condición resulta de la unidimensionalidad del constructo y la segunda del tamaño de la población referente para establecer la invarianza (Linn, 1990; Osterlind, 1992).

Un elemento adicional, que guía la selección de Indicadores tanto para el Análisis de Ítems, como para la Fiabilidad, es el tipo de interpretación de las puntuaciones que se requiere. Así, la dicotomía Normativa-Criterial está a la base de esta selección.

Las pruebas estandarizadas de Rendimiento como Indicadores de Resultados no son interpretadas individualmente, por lo que, en principio, no sería necesario ningún elemento que coadyuve a la interpretación. Así, respecto a la selección de indicadores,

¹⁰ Y debe ser refrendada por la acumulación de evidencias empíricas, tanto desde una base de investigación experimental como correlacional.

el Análisis de parámetros de los ítems y el sesgo, junto a comprobaciones de la fiabilidad como consistencia global serían suficientes para un planteamiento métrico adecuado de las pruebas.

Sin embargo, el contexto de uso de estas pruebas —aunque no requieran de un Estándar para la interpretación individual de puntuaciones— lleva a que se necesite algún procedimiento global. De este modo, estas pruebas se utilizan en Evaluaciones sobre la Calidad de un Sistema, un Centro o un Programa y, por ello, debe tenerse en cuenta que de no acompañarse de ningún elemento interpretativo la Evaluación quedará en el terreno meramente descriptivo.

En éste ámbito, las informaciones normativas son indudablemente la base de análisis y la referencia más clara respecto a las características del Dominio evaluado. No obstante, es conveniente que el Comité de Expertos que desarrolla el Análisis y Especificación del Dominio establezca además un Estándar¹¹ —basado en juicio— que identifique, al menos, los niveles mínimos de competencia aceptables como indicador de suficiencia del sistema. El Estándar operativiza la idea de calidad. Este tipo de Estándares, son especialmente útiles en estudios Longitudinales, pues es conocido que cuando trabajamos con grandes muestras, pequeñas diferencias se identifican como diferencias estadísticamente significativas, y es necesario algún tipo de criterio que nos permita valorar la “cualidad de la diferencia”.

Para el desarrollo de este tipo de Estándares se puede trabajar desde Metodologías específicas de Estándares basados en los ítems. Son especialmente de interés para este tipo de pruebas los procedimientos desarrollados para situaciones multivariadas de decisión, como por ejemplo las propuestas de modificación del método de Angoff (Hambleton y Plake, 1995), el procedimiento de Jaeger (1993) o la síntesis formulada por Putnam, Pence y Jaeger (1995). Estos procedimientos abordan la toma de decisiones en tareas complejas, partiendo de la base de que lo que se busca es un perfil de ejecución a través de un conjunto de dimensiones relevantes; lo que es el caso de una buena parte de las situaciones que se encuentran dentro de este tipo de pruebas. En definitiva, constituyen un avance en la línea de operativizar la idea de calidad a partir de los contenidos evaluados, reteniendo la complejidad consustancial a la magnitud del Dominio a que se refieren este tipo de pruebas.

• *Pruebas Estandarizadas de Certificación y de Admisión.*

Estas pruebas tienen por objeto recoger la información que permita certificar que una persona ha superado administrativamente un determinado nivel educativo o que tiene los conocimientos necesarios para ser admitido en un programa de formación de amplio espectro, por ejemplo, en la enseñanza universitaria¹². Son pruebas que, por

11 Entendemos por Estándar la puntuación en el Dominio que indica el nivel mínimo de competencia. Se trata de la expresión de este nivel en la Escala de puntajes verdaderos, teóricos, libres de error (Jornet y Suárez, 1989b).

12 En nuestro país existen pocas experiencias aún desarrolladas con este tipo de pruebas. Estudios de interés a este respecto son los de Tourón (1985) y Toca y Tourón (1989), en el ámbito universitario, o en relación a la Educación General Básica los trabajos llevados a cabo por Rivas et al. (1986), que desarrollaron una línea de investigación que condujo a la elaboración de pruebas para los finales de Ciclo en la EGB en diversas materias.

tanto, se orientan a un universo instruccional muy amplio, cuya definición es básicamente empírica —es decir, muy operativizada—, y que se centran muy especialmente en el producto educativo.

Así, si se pretende establecer un nivel generalizado que certifique unos conocimientos mínimos para superar la Secundaria Obligatoria, obviamente estamos hablando de abordar la medición y evaluación de un Dominio educativo que se extiende a lo largo de cuatro años y que está concretado en un número importante de materias diferenciadas. Además, parece razonable tener presente que para esta situación no existen aportaciones teóricas que nos permitan extraer conclusiones ciñéndonos a la valoración de unas cuantas dimensiones. Como mucho, podremos efectuar una definición bastante pormenorizada del Universo de Medida eligiendo algún punto de referencia como pueden ser los textos legales que reflejan las orientaciones y objetivos necesarios en estos niveles. En relación directa con esta cuestión, si no se dispone de un marco teórico sólido de referencia y se debe abarcar un universo muy amplio es básicamente imposible abordar una evaluación del proceso, por lo que estas pruebas se suelen concentrar en la valoración del producto educativo. Las reglas de conexión entre el Universo de Medida de referencia y la prueba concreta no se pueden especificar de forma exhaustiva, por los mismos motivos que acabamos de apuntar.

Las referencias a la definición del Dominio Educativo, que señalamos para las pruebas anteriores, son aplicables aquí. Únicamente debe tenerse en cuenta que en la selección de unidades del Dominio prevalecerán los juicios acerca de la relevancia de los ítems como expresión de competencia, dado que ello es especialmente importante para poder establecer el Estándar.

• *¿Qué tipo de interpretación se requiere en estas pruebas?*

A diferencia de las pruebas descritas anteriormente, hay que considerar que a partir de estas pruebas se pretende realizar una interpretación específica del nivel de competencia de cada persona, por lo que el planteamiento global de desarrollo de las pruebas varía sustancialmente en la selección de indicadores. Precisamente debido a esta referencia individual en el objetivo de valoración, en este contexto, cuando se trata de la valoración de personas pertenecientes a una población muy amplia resulta especialmente importante enfatizar el análisis del sesgo para asegurar la equidad de la evaluación.

El establecimiento del estándar de superación es difícil que se refiera exclusivamente a un valor absoluto. En este sentido, hay que tener presente que la propia amplitud y heterogeneidad del contenido hacen muy difícil poder definir exactamente cuál es el nivel mínimo exigible —mediante objetivos o conocimientos específicos— a un sujeto para alcanzar el nivel de competencia. Además, estas pruebas afectan al conjunto de la sociedad y, por ello, deben participar muy directamente en este proceso de decisión los diversos colectivos implicados. Así, es conveniente llevar a cabo un proceso de determinación del estándar de tipo mixto, integrando los criterios absolutos con las consecuencias que de su aplicación se derivarían para diferentes colectivos.

De este modo, es muy importante tener en cuenta en este trabajo la minimización de los Errores de Selección, por lo que las labores de adecuación del Estándar como

Punto de Corte¹³ son especialmente relevantes. La determinación de la puntuación de corte debe desarrollarse a partir de un proceso iterativo en el que se conjuguen técnicas de juicio con análisis empíricos y en el que el estudio de las consecuencias de aplicación del Estándar moderen las aplicaciones de juicio. En este contexto es básica la retroalimentación de información al Comité de Expertos que desarrolla las pruebas.

En estos procedimientos se tiende a seleccionar la puntuación de corte como aquella que maximiza la fiabilidad y minimiza los Errores de Selección (tipo I y tipo II). No obstante, atendiendo a la Razón de Pase¹⁴, si es que está prefijada, el Comité de Expertos puede valorar la utilidad diferencial de asumir decisiones con Error tipo I o tipo II, de forma que se integren en la determinación de la puntuación de corte la composición de aquellos errores evaluativos que resulte menos lesiva para el adecuado uso del estándar.

Como en el caso de las pruebas anteriores estas pruebas deben sustentar su desarrollo como Modelo de Medida sobre la Teoría de Respuesta al Ítem. Sin embargo, en la selección de indicadores para el análisis de ítems y para la fiabilidad es preciso tener en cuenta la existencia del estándar criterial. Por ello, en la determinación de la fiabilidad deben contemplarse indicadores de Consistencia de la Decisión, de forma que pueda valorarse la capacidad de la prueba para diferenciar, al menos, entre sujetos que tienen y no tiene el nivel mínimo de competencia en el Dominio de Referencia. Por su parte, en el análisis de ítems se atenderá especialmente a la identificación de indicadores de discriminación, en los que deberán incluirse formulaciones que tengan en cuenta —además de la capacidad global de discriminación— la actuación consecuente con el Punto de Corte fijado.

• *Pruebas de Nivel y de Dominio.*

Las pruebas de Nivel y las de Dominio las podemos considerar como variaciones de las anteriores, diferenciándose en virtud de la amplitud del Dominio Educativo a que se refieren —que es más reducido— o bien en relación a la amplitud de la población a la que van dirigidas —que también suele ser más específica—. Tienen en parte por tanto unos objetivos y características semejantes a las anteriores. En este caso, nos estamos refiriendo a pruebas que aporten información, por ejemplo, sobre si un estudiante ha alcanzado el nivel suficiente como para pasar de un curso a otro en una materia, o bien si ha superado los niveles mínimos exigidos en un programa de formación concreto¹⁵ —por ejemplo, un programa de reciclaje para docentes sobre técnicas de observación en el aula—. El hecho de referirse a un programa educativo mucho más concreto conlleva diferencias sustanciales que se pueden sintetizar en las siguientes:

13 Entendemos por Punto de Corte aquella puntuación en la prueba que expresa el nivel mínimo de competencia. Proviene del Estándar y constituye el ajuste empírico del mismo, teniendo en cuenta criterios de fiabilidad (Jornet y Suárez, 1989b).

14 % de sujetos que pueden ser admitidos, por ejemplo, en un programa.

15 Se ha informado de algunos desarrollos específicos de pruebas —insertas en Modelos evaluativos— que podrían ser identificables en esta categoría como los trabajos de Rodríguez Lajo (1986), Jornet (1987) y Jornet et al. (1993).

- a) se puede dar una definición del universo de referencia tanto empírica como teórica — esto último especialmente en los programa más concretos—,
- b) se puede aspirar a la valoración del proceso y no sólo del producto. Al ser pruebas referidas a dominios más concretos cabe identificar unidades en el Dominio, a partir de las cuales se puedan realizar inferencias acerca de los procesos,
- c) las reglas de conexión entre el universo y la prueba están mucho más determinadas —en numerosas ocasiones completamente determinadas—,
- d) en el análisis de ítems, junto a los indicadores de los parámetros básicos resulta de interés (al ser pruebas de nivel o de evaluación de un programa específico) la sensibilidad instruccional, como expresión de la capacidad de los ítems para discriminar las adquisiciones propias del programa,
- e) la amplitud de las poblaciones de referencia puede condicionar el Modelo de Medida adecuado a cada caso. Así, se debe distinguir entre las pruebas que se orientan a poblaciones amplias y las que se destinan a ámbitos más concretos. Por ejemplo, hay que diferenciar entre una prueba de Cálculo para primero de Primaria aplicable a estudiantes valencianos y otra destinada a evaluar la competencia alcanzada en un programa de formación para la participación de Equipos Directivos de Centros. El Dominio se concreta en ambos caso, pero la primera situación va dirigida a una población amplia y le son aplicables los mismos referentes de Medida que los ya comentados en los casos anteriores, mientras que en la segunda situación las pruebas se deberán sustentar en la Teoría Clásica del Test y en indicadores que provienen del ámbito de la Evaluación Referida al Criterio¹⁶.
- f) el estándar de referencia está normalmente basado en un criterio absoluto —aunque se den, obviamente, casos en que se utiliza una combinación con información normativa—.

En cualquier caso, este tipo de pruebas se sitúa entre las pruebas amplio espectro y las de Aula, adoptando características que les son propias a los dos enfoques.

PRUEBAS DE CLASE O DE USO EN EL AULA

Las pruebas de Clase o pruebas de Aula hacen referencia a las que puede utilizar el profesor para la evaluación de sus alumnos. No obstante, es en este ámbito donde probablemente se aprecia una peor aceptación de las pruebas estandarizadas. Y ello, porque se atribuye a la Estandarización condiciones que alejan estas pruebas de la individualización.

Sin embargo, debe tenerse en cuenta que ambos —Estandarización e Individualización— no son conceptos necesariamente contradictorios; más bien, es el tipo de uso que se realice de las pruebas lo que puede enfrentarlos. La Estandarización mejora

¹⁶ En ambos casos se pueden utilizar aportaciones derivadas de la Teoría de la Generalizabilidad para conseguir indicadores más consistentes de los parámetros de la prueba, especialmente en aquellos casos en los que no sea adecuado utilizar los modelos TRI (Brennan, 1983; Shavelson y Webb, 1991).

esencialmente las condiciones de objetivación de la medida. Y ello a veces se olvida por parte de los detractores de las pruebas, poniendo el énfasis sobre aspectos propios de la individualización que, sin embargo, podrían ser atendidos con pruebas estandarizadas, sin considerar que los procedimientos no-estandarizados no resuelven adecuadamente los problemas derivados de la subjetividad del observador o el evaluador.

Por otra parte, el marco derivado de la LOGSE ha puesto de manifiesto la necesidad de individualizar o personalizar los diseños curriculares. Un problema práctico al que habitualmente aluden los profesores es la falta de metodología adecuada para el desarrollo de las Adaptaciones Curriculares Individualizadas. Obviamente, las soluciones no están sólo en la Medida y la Evaluación, pero éstas constituyen un elemento instrumental inicial desde el que abordar la solución de este problema.

Así, debe tenerse en cuenta que el desarrollo de pruebas estandarizadas para la evaluación de una unidad didáctica es probablemente el marco donde puede disponerse de pruebas más válidas. Esto es así, dado que los Dominios Educativos de referencia en estos casos constituyen Universos de Medida finitos, claramente especificables, concretos.

Este hecho afecta directamente a la Validez de Contenido, pudiendo aspirar en este contexto a pruebas más representativas del Dominio Educativo del que se derivan. Por otra parte, si se estructura el Dominio Educativo de forma perfectamente asociada al planteamiento metodológico-didáctico, la Validez de Constructo también puede verse beneficiada. Esto es así, no sólo por el hecho de la asociación trabajo de aula-sistema de evaluación (que sería una expresión más concreta de la Validez de Constructo, como Validez Curricular), sino muy especialmente por las características del desarrollo del aprendizaje, en el que se podrá reflejar el constructo teórico que esté a la base del diseño curricular y del enfoque metodológico-didáctico del programa.

Además, los Dominios Educativos, en estos casos, se refieren a unidades didácticas —o lecciones— por lo que incluyen pocas unidades, lo que favorece el micro-análisis de todas las tareas-ítems implicados en el Dominio. Pueden permitir, pues, una definición exhaustiva de la población de conductas que pongan de manifiesto las adquisiciones (habilidades, capacidades y destrezas) a que hace referencia un Dominio. De esta forma, en la definición de este tipo de Dominios cabe identificar unidades de medida, a partir de las cuales se puedan inferir interpretaciones procesuales bastante precisas, así como de productos específicos.

Ello favorece que este tipo de pruebas pueda estar muy bien adaptado para la medición y evaluación de procesos y productos de aprendizaje¹⁷. De este modo, las posibilidades de interpretación se abren: no sólo se puede interpretar un nivel de competencia —a partir de la puntuación total— sino explicar el nivel, informando de los procesos de adquisición —a partir de la interpretación particular de los ítems—.

Este hecho se ve favorecido porque la situación de Medida que puede darse en un Aula no tiene por qué limitarse a una situación de examen habitual tipo test. En este

17 Recientemente se ha venido informando de propuestas de interés en nuestro ámbito educativo, como las de Buendía y Salmerón (1994) o las de Toboso (1995 a y b).

marco, pueden formar parte de la “prueba” diversos tipos de elementos: desde ítems clásicos de lápiz y papel hasta ítems micro-situacionales en los que la valoración provenga de la observación de la tarea que realiza cada individuo.

La definición del Dominio, establecimiento de Reglas de generación de ítems y escritura de ítems, la realiza —como en casos anteriores— un comité de expertos, pero en esta situación, está compuesta por los profesores de una materia (o departamento, o equipo de ciclo).

En el trabajo a realizar por el Comité hay que tener en cuenta los siguientes aspectos:

- Antes de desarrollar la prueba es esencial definir el rol que ésta tendrá dentro de los recursos evaluativos de que disponga el profesor. El contexto ideal de uso es aquel en el que se identifican fuentes múltiples y diversos instrumentos.
- En el desarrollo del Análisis del Dominio la reflexión deberá orientarse hacia elementos de relevancia de los ítems respecto de los objetivos que pretenden medir (Congruencia Ítem-Objetivo) así como respecto a la representatividad de los ítems —como situación evaluativa— en relación a los planteamientos metodológico-didácticos seguidos en el Aula.
- Aunque en este contexto no resulta tan trascendente el análisis del Sesgo de los ítems, su anticipación —por procesos de juicio— es una labor de especial interés. Así, junto a variables básicas como Sexo o Lengua —en Comunidades bilingües—, dependiendo del nivel educativo, pueden producirse sesgos en situaciones de apertura del currículum. De esta manera, pueden haber alumnos reforzados en su aprendizaje de una materia por el efecto del aprendizaje de otras opcionales. Este tipo de sesgo sería importante identificarlo a efectos de determinar adecuadamente el nivel de competencia a que puede aspirarse en la materia en la que se desarrolla la prueba.
- No obstante, aunque existan sesgos de los que necesariamente tengamos que independizar a las pruebas, otros —como el último mencionado— puede ser difícil de eliminar, por lo que al menos es importante identificarlos, conocerlos, y asignarles valor diagnóstico o modulador de las decisiones evaluativas.
- El análisis del sesgo en este caso se basa más sobre procesos cualitativos de juicio que sobre la comprobación empírica de los resultados obtenidos por las personas evaluadas. El factor clave para poder llevar a cabo una comprobación adecuada es el tamaño de la muestra, en estos casos muy reducida.

Respecto a los componentes técnicos derivados del Modelo de Medida, en este contexto no puede utilizarse la Teoría de Respuesta al Ítem, dado que el tamaño de los grupos que se trabaja es muy reducido.

No obstante, dentro de la Teoría Clásica de los Tests se dispone de indicadores suficientes que pueden, realizando las adecuadas adaptaciones en su uso e interpretación, operacionalizar los análisis necesarios. Asimismo, en el marco de la Evaluación Referida al Criterio existen múltiples indicadores de fácil utilización que racionalizan la lógica de selección criterial.

Sin embargo, hay que tener presente que las características derivadas del objeto de

medida y del uso de la prueba, así como las que devienen de los tipos de distribución que podemos encontrar en pequeñas muestras¹⁸, conllevan necesarias modificaciones en la utilización e interpretación de indicadores clásicos. De esta manera, características básicas empíricas útiles en la selección de ítems del Modelo Clásico no serían deseables aquí. El valor de los parámetros e indicadores radica no tanto como elemento de selección de los ítems (que se sustentará preferentemente sobre procesos de juicio) como elemento de información para el grupo de profesores —Comité— que desarrolla la prueba. Su valor como indicadores de selección de ítems, se circunscribe a la comprobación de las hipótesis funcionales que el Comité haya anticipado para los ítems, respecto a la dificultad teórica y su capacidad de discriminación primordialmente (Jornet y Suárez, 1994 ; Rivas, Jornet y Suárez, 1995).

¿Qué tipo de interpretaciones se requieren?

En este contexto, en donde se deben tomar decisiones acerca de la promoción de cada persona en su aprendizaje, es preciso conocer adecuadamente su posición respecto del Dominio Educativo, por lo que la interpretación necesariamente debe ser criterial, basada en un estándar absoluto.

Para el desarrollo del Estándar y su especificación como puntuación de corte la aportación del Comité de Expertos es nuevamente esencial. Entre los métodos en que pueden apoyarse, en este contexto de desarrollo de pruebas, cobran especial relevancia aquéllos que utilizan como información para retroalimentar al Comité en su proceso de determinación, el análisis de las consecuencias de su aplicación sobre sujetos conocidos (Livingston y Zieky, 1982). Estos usos, en la práctica, se convierten en evidencias de validación de la puntuación de corte.

En cualquier caso, los procesos de evaluación en el Aula hay que considerar que no se debe aspirar a que se sustenten sobre pruebas estandarizadas exclusivamente. La oportunidad de su utilización depende fundamentalmente de que sean adecuadas al tipo de materia que se pretenda evaluar. La defensa del uso de este tipo de pruebas para la evaluación debe realizarse desde el marco en que realmente sean más útiles (fiabes y válidas) que otras alternativas o técnicas evaluativas. Asimismo, debe tenerse en cuenta que las pruebas estandarizadas deben derivarse desde un programa educativo bien establecido y, como instrumento están al servicio del mismo. Un peligro genérico que nace del uso de cualquier sistema de evaluación es que acaben siendo los instrumentos los que constituyan la referencia para el desarrollo del programa, acabando por condicionar su uso (De la Orden, 1993).

PRUEBAS DE PROPÓSITO DIAGNÓSTICO

Las *Pruebas de Propósito Diagnóstico* rompen con el discurso de lo más general a lo más concreto que relaciona los tres tipos que hemos expuesto. Así, mientras en las pruebas revisadas con anterioridad el objetivo es, en términos generales, valorar las

18 No sería aceptable anticipar como efecto educativo que el Rendimiento se distribuirá como la Curva Normal. Es más lo habitual —y deseable— sería que los efectos educativos fueran asimilables a distribuciones beta, con tendencia asimétrica negativa.

adquisiciones sobre un programa educativo, más o menos amplio, aquí se trata de poder indagar respecto a la posible existencia de determinados problemas de aprendizaje y cuáles son sus características concretas. El objetivo es, pues, delimitar si se da un determinado problema en el proceso normal de adquisiciones que desarrolla un sujeto y poder extraer información sobre la cualidad de tal problema, de modo que se pueda orientar mejor la subsiguiente intervención.

En este grupo incluimos dos grandes tipos de pruebas: referidas al Currículum y de diagnóstico propiamente dichas.

Las pruebas estandarizadas referidas a un currículum tienen por objeto indagar acerca de la posición de un sujeto respecto a un Diseño Curricular dado (que actúa como Dominio Educativo). Las hemos clasificado aquí porque su propósito generalmente es de tipo diagnóstico, dado que se trata de recabar información independiente de las calificaciones escolares —y sin finalidad de uso en el contexto del Aula— acerca de si los sujetos tienen adquisiciones básicas correspondientes con su desarrollo curricular-escolar, o bien presentan disfunciones¹⁹.

Generalmente, estas pruebas pretenden abarcar Dominios amplios, correspondientes a dimensiones que se identificables a través de un Diseño Curricular de largo alcance (como por ejemplo, Numeración, Cálculo Mental, Interpretación de Datos o Resolución de Problemas). Sin embargo, en el análisis del Dominio prevalece la identificación de los elementos clave que se asocian a diferentes etapas de adquisición. Esta identificación es la que permite situar al sujeto en su nivel de aprendizaje en el Dominio.

Por su parte, las pruebas Diagnósticas propiamente dichas tienen como finalidad no sólo determinar la posición del sujeto en el Dominio de referencia, sino describir adecuadamente los elementos deficitarios con el fin de planificar la intervención. En este sentido, las unidades del Dominio deben estar claramente definidas y previamente analizadas respecto a su asociación con unidades de intervención.

Es habitual que en estas situaciones se parta de un marco teórico que define dimensiones respecto de las cuales se puede concentrar la información significativa para la toma de decisiones, aunque no siempre tenga la consistencia teórico-metodológica debida (De la Orden et al. 1994).

Como señala Oosterhof (1994), las pruebas de propósito diagnóstico deben ser utilizadas con cautela, pues la investigación básica acerca de los constructos en ellas implicados, todavía es escasa²⁰.

En relación directa con esto, el objetivo de la evaluación con estas pruebas es mixto, en el sentido que pueden estar orientadas al proceso educativo, al producto o a cualquier combinación de ambos objetivos. Esto conlleva que las pruebas estén basa-

19 Aunque no son muy frecuentes, existen ya algunas pruebas de interés como por ejemplo la Escala Key-Math R de Connolly (1988), adaptada por Marí (1996) a nuestro contexto educativo o algunas otras desarrolladas directamente en el mismo, como la Batería de Pruebas de Lenguaje FCI (Bartolomé, et al., 1985).

20 Sin embargo, se pueden identificar ya desarrollos muy adecuados (como la prueba CRIL de Lenguaje de Wiig, 1990, de la que parte el desarrollo de la prueba ICL de Puyuelo y Renom —1993— y Puyuelo, Renom y Solanas —1995—).

das en unas reglas de conexión con el universo tan específicas como lo permita las características y la amplitud del mismo. Finalmente, el estándar en que se basa la decisión suele ser una combinación de indicadores absolutos y normativos. Esto es así dado que para la determinación de la existencia de un problema suele ser tan útil emplear definiciones absolutas que reflejen las claves de su identificación (nivel en que se produce un problema, patrón procesual del mismo, etc...) como la información relativa al grupo de pertenencia para situar la dimensión característica del mismo —por ejemplo, un problema de inversiones en la lectura dependerá tanto de una determinada frecuencia concreta como de la situación relativa dependiendo del grupo de edad al que pertenece el sujeto—.

PRUEBAS INDIVIDUALIZADAS

En este caso el objeto es proporcionar un sistema de recogida de información muy flexible que se ajuste a las características de cada sujeto o situación de medida y que proporcione, por ello, una información más rica y significativa en los puntos críticos. Como se aprecia en el Cuadro 2, no siempre la unidad de referencia es un sujeto concreto, pudiendo ser un currículum o programa completo. Además, la adaptación puede realizarse de forma estática o dinámica. En el primer caso la prueba entera se construye en función de las características o directrices del grupo o situación de referencia, mientras que en el segundo caso es el propio rendimiento el que proporciona el patrón de referencia para la adaptación sucesiva de la prueba.

Estamos hablando de pruebas que precisan de una definición lo más exhaustiva posible del Dominio Instruccional de referencia y de unas reglas sumamente concretas de relación entre el Universo Instruccional y la prueba. De no producirse estas condiciones no se podría establecer bien el ajuste para cada situación o individuo.

En general, las pruebas adaptadas se fundamentan en bancos de objetivos e ítems que ofrecen algunas grandes organizaciones públicas y privadas. Los dos formatos esenciales se dan en función de que sea la propia empresa u organización que facilite la adaptación ya completada al usuario final o que se le suministre la información y una herramienta informática a este último para que pueda hacer esta adaptación por sus propios medios. Así, los sistemas AIMS (Academic Instructional Measurement Systems) de The Psychological Corporation, ORBIT (Objective-Referenced Bank of Items and Tests) de CTB/McGraw-Hill o MULTISCORE de The Riverside Publishing Company están compuestos por unos centenares a miles de objetivos y muchos miles de ítems que abarcan la mayor parte de los ámbitos curriculares de la educación primaria y secundaria.

La adaptabilidad en el caso de situaciones o programas dependerá de la cantidad de opciones disponibles sobre el dominio (objetivos, ítems, etc.), de las informaciones sobre la estructura y características métricas de los elementos (dificultad, discriminación, elección de alternativas, recomendaciones asociadas, etc.) y de la existencia de mecanismos para integrar variaciones propias de cada situación en la prueba (herramientas que permitan el desarrollo de objetivos, ítems con diversas variantes, recomendaciones, etc.).

En el caso de los tests con adaptación instantánea a la ejecución por el sujeto —como el MicroCAT Testing System (Assessment Systems Corporation, 1988) o los WICAT Systems (1986)— sirven esencialmente las ideas que acabamos de apuntar. No obstante, la adaptabilidad en estos casos se incrementa cuando estas pruebas se pueden adaptar mejor al sujeto en la situación específica de aplicación que sirve como referencia (sea el programa, el sistema específico de recuperación, etc.). Asimismo, la información que se puede obtener en estas situaciones es tan rica que difícilmente se puede encarar una prueba de este tipo solamente en función de una valoración del producto. No obstante, este es el enfoque predominante todavía, pues se ha producido una adaptación excesivamente mimética respecto a las pruebas tradicionales. Piénsese que estamos hablando de pruebas que por su complejidad, normalmente, precisan de un soporte de tecnología informática, ya que el único medio que facilita una gran precisión y variedad en la recolección de información. De hecho, este tipo de medida es el horizonte natural de los sistemas EAO al incorporar la evaluación dinámica que se precisa en estos casos. Hoy en día, con todo, la mayoría de los sistemas EAO están lejos de adoptar las posibilidades de medida inherentes a este tipo de pruebas —de hecho, tienen serios problemas para cubrir las mínimas directrices que garanticen una valoración del rendimiento—. Existen, no obstante, algunas excepciones que constituyen caminos muy sugerentes, como la experiencia del Cognition Technology Group en la Universidad de Vanderbilt (Goldman, Pellegrino y Bransford, 1994), con planteamientos integrales de enseñanza y evaluación-medida que pueden aportar respuestas a algunas necesidades de transformación que ya hemos señalado.

La calidad de la Definición del Dominio de estas pruebas basadas en bancos de ítems es bastante elevada²¹. En cualquier caso, el nivel de especificación que requieren permite una valoración ajustada y actualizada de estos sistemas que sirva de referencia para nuestra actuación. Por ejemplo, se realizan revisiones de estos sistemas periódicamente que nos añaden referentes de validez y utilidad de los mismos (Naccarato, 1988).

Los Modelos de Medida asociados a las aplicaciones más consistentes están basados en Teoría de Respuesta al Ítem, tanto para la composición del banco de referencia como para su gestión en cada situación de evaluación concreta (Kingsbury y Zara, 1989). No obstante, como ya hemos señalado, el campo es muy heterogéneo y desestructurado, hallándose incluso pruebas que no están soportadas por modelo alguno de medida. En definitiva, para los proyectos de una cierta envergadura —respaldados por instituciones o empresas de suficiente solvencia— las herramientas disponibles en los modelos TRI constituyen la recomendación universalmente aceptada. Las limitaciones son las mismas que ya se han descrito respecto a otras pruebas y las ventajas son aún mayores, al entramarse los procedimientos con mayor facilidad en una estructura de aplicación basada en ordenador. Queda por resolver, a pesar de todo, una amplia variedad de temáticas y situaciones para las que, al igual que en otros contextos de evaluación, no cesan de proponerse alternativas parciales que, siendo muy

21 Otra cuestión bien diferente la constituyen los sistemas de evaluación ligados a las aplicaciones EAO, cuya calidad metodológica general es muy desigual.

CUADRO 2
SÍNTESIS DE PROCEDIMIENTOS DE PRUEBAS INDIVIDUALIZADAS

Tipos de Prueba o procedimiento	Objetivo	Características de las Tareas/Pruebas
PRUEBAS DE ADMINISTRACIÓN INDIVIDUAL	Mejorar la precisión en la estimación de la habilidad del sujeto	Las tareas están graduadas, en ocasiones se basa la administración en senderos de decisión.
PRUEBAS DE FORMAS MÚLTIPLES	Obtener múltiples medidas paralelas, equivalentes —o al menos comparables— de cada sujeto.	Pruebas estadísticamente paralelas
PRUEBAS ESTANDARIZADAS DE NIVELES MÚLTIPLES	Simplificar la medida ajustándola al nivel de habilidad del sujeto	Ítems basales; determinaciones del nivel inicial de partida de la prueba.
TESTS ADAPTATIVOS • DE NIVEL DIAGNÓSTICOS	Maximizar la precisión de la estimación de la habilidad de cada sujeto, utilizando el mínimo número de ítems Diagnosticar las dificultades de aprendizaje concretas del sujeto	Ítems clave-característicos de niveles. Selección específica de los ítems a administrar en función del nivel inicial demostrado por el sujeto en ítems de prueba.
SOPORTES TECNOLÓGICOS A LA INDIVIDUALIZACIÓN • BANCOS DE ÍTEMS TRADICIONALES	Automatizar la construcción de pruebas a partir de una definición genérica del Dominio, con ítems cerrados e identificados por sus parámetros.	Pruebas ajustadas a diseños curriculares y adaptadas a niveles específicos.
• BANCOS DE REGLAS DE GENERACIÓN DE ÍTEMS	Automatizar la construcción de ítems, y en algunos casos, incluso su administración. Generar múltiples pruebas paralelas.	Tests paralelos en contenido, diferentes para cada sujeto y que pueden ajustarse a los diversos curricula.

sugerentes, no se han estructurado en ningún planteamiento suficientemente sólido hasta el momento presente.

Por lo que respecta a los estándares, aunque también utilizan información respecto a criterios diferenciales, lo fundamental de las decisiones consiste en las definiciones absolutas que incorporan. De hecho, en muchos casos los referentes diferenciales son parciales o imposibles, en función de la adaptación que se realice —por ejemplo, si se añade una cantidad sustancial de modificaciones en la definición de algún subdominio educativo todo referente normativo a este respecto carecería de validez—. En otras

situaciones, la información normativa solamente puede actuar como referente marco relativamente alejado, dado que el propósito de estas pruebas suele ser más formativo u orientado a la recuperación. Desafortunadamente existe poco trabajo hecho en este ámbito e incluso las orientaciones reflejadas a este respecto en las "normas y orientaciones de actuación" (APA, 1986) no parecen haber madurado suficientemente la necesaria adaptación a estas situaciones de medida-evaluación.

ALGUNAS NOTAS FINALES

La evolución de los métodos de construcción de pruebas aporta una base bastante sólida para el desarrollo de instrumentos estandarizados de evaluación. Para nuestro ámbito educativo, los desarrollos son muy escasos, aunque crecientes, en consonancia con la progresiva implantación de actividades de evaluación. El arraigo de una cultura evaluativa sin duda conllevará la necesidad de utilizar instrumentos mucho mejor contruidos y adaptados que los que habitualmente se utilizan. La demanda de calidad también afectará a los instrumentos evaluativos. Sólo será posible responder a estos retos si abordamos decididamente el desarrollo de instrumentos de medida educativa, que respondan a las necesidades de los diversos programas y fenómenos a evaluar.

La institucionalización de la revisión del sistema educativo y de los diferentes componentes y actores del mismo es una realidad a la que necesariamente se debe responder con instrumentos mejor diseñados. Incluso, hechos evaluativos que afectan muy directamente a nuestra sociedad —como es la selectividad universitaria— en la actualidad aún se desarrolla sobre esquemas imprecisos, que hacen que ésta no responda en definitiva al sistema y que no se pueda hablar de equidad evaluativa. Actuaciones como la selectividad están reclamando respuestas profesionales evaluativas que, al menos, integren las opciones metodológicas disponibles.

En el campo del Diagnóstico Educativo también es evidente la carencia de instrumentos. De hecho no se dispone de Baterías a lo largo de los diferentes niveles y áreas educativas que cubran las dimensiones esenciales del Rendimiento. Otro tanto puede decirse respecto a las versiones individualizadas, como las Pruebas Asistidas por Ordenador, que además deben reivindicar su existencia frente a los exiguos sistemas de evaluación que incorporan las aplicaciones de Enseñanza Asistida por Ordenador.

Respecto a las evaluaciones en el Aula deben estar soportadas sobre una profunda reflexión por parte de los profesores acerca de los componentes de sus diseños curriculares. Un aspecto central de esta reflexión debe ser el sistema de evaluación. El desarrollo de instrumentos no tiene por qué ser la finalidad, pero sí constituye un buen medio de revisión de los componentes de un programa educativo, así como de los elementos que inciden en su realización. Incorporar elementos propios de las pruebas estandarizadas, como son el análisis de los Dominios Educativos o el Desarrollo de Estándares, aunque no se persiga ni se llegue a una estandarización completa, supone indudablemente integrar elementos de mejora de los procesos evaluativos. Obviamente, en muchas ocasiones se afirma que el profesorado no ha sido formado para abordar este tipo de procesos. La respuesta es clara: es necesario reforzar la

formación de estos profesionales en estas áreas, dado que son instrumentales para su actuación docente.

Por último, la estandarización de la medida, como base para la evaluación, si quiera en sus versiones más "tecnológicas" —como las Pruebas Asistidas por Ordenador—, no implica necesariamente un empobrecimiento de la información necesaria en la evaluación. Existen suficientes vías de trabajo para incorporar definitivamente la medida de tareas complejas, el proceso de construcción de los conocimientos, el aprendizaje cooperativo o el pensamiento crítico. En buena medida, nos tenemos que comprometer en realizar un esfuerzo por avanzar en esas direcciones y tratar de situarnos en línea con los países más desarrollados en estos ámbitos disciplinares.

REFERENCIAS BIBLIOGRÁFICAS

- AMERICAN PSYCHOLOGICAL ASSOCIATION (1986) *Guidelines for computer-based tests and interpretation*. Washington, D.C.: Autor.
- ANGOFF, W.H. (1988) Validity: An evolving concept. En H. WAINER y H.I. BRAUN (Eds.) *Test Validity*. Hillsdale, NJ: LEA.
- BARTOLOMÉ, M.; BISQUERRA, R.; CABRERA, F.; ESPÍN, J.V.; MATEO, J. Y RODRÍGUEZ, L.I. (1985) *Batería de Pruebas de Lenguaje Final de Ciclo Inicial*. Barcelona: CEAC.
- BRENNAN, R.L. (1983) *Elements of Generalizability Theory*. Iowa City, IA: American College Testing Program.
- BUENDÍA, L. y SALMERÓN, H. (1994) Construcción de pruebas criterioles de aula. *Revista de Investigación Educativa*, 23, 405-410.
- CERI/INES (1995) *Education at a Glance*. OECD Indicators. París: OECD.
- CONNOLLY, A.J. (1988) *Key Math Revised: a diagnostic inventory of essential mathematics*. Circle Pines, Minnesota: American Guidance Service.
- DE LA ORDEN, A. (1993) Influencia de la evaluación del aprendizaje en la eficacia de la enseñanza. *Revista de Investigación Educativa*, 22, 7-42.
- DE LA ORDEN, A.; GAVIRIA, J.L.; FUENTES, A. y LÁZARO, A. (1994) Modelos de construcción y validación de instrumentos diagnósticos. *Revista de Investigación Educativa*, 23, 129-178.
- DE MIGUEL, M. et al. (1994) *Evaluación para la calidad de los Institutos de Educación Secundaria*. Madrid: Escuela Española.
- GIL, G.; GONZÁLEZ, A. y SUÁREZ, J.C. (1995) Un modelo de construcción de pruebas de rendimiento para la evaluación de las enseñanzas mínimas en la Educación Primaria. En AIDIPE (Comp.) *Estudios de Investigación Educativa en Intervención Psicopedagógica*. Valencia: AIDIPE.
- GOLDMAN, S.R., PELLEGRINO, J.W. y BRANSFORD, J.D. (1994) Assessing programs that invite thinking. En E. BAKER y H.F. O'NEIL Jr. (Eds.) (1994) *Technology Assessment in Education and Training*. Hillsdale, NJ: LEA.
- HALADYNA, T.M. (1992) Context dependent ítem sets. *Educational Measurement: Issues and Practice*, 11, 11-25.
- HAMBLETON, R.K. (1984) Validating the tests scores. En R. BERK (De.) *A guide to*

- Criterion-Referenced Tests construction*. Baltimore, Mass.: Johns Hopkins University Press.
- HAMBLETON, R.K.; SWAMINATHAN, H. (1985) *Ítem Response Theory: Principles and Applications*. Norwell, MA: Kluwer.
- INCE (1996) *Lo que aprenden los alumnos de 12 años. Evaluación de la Educación Primaria. Datos básicos*. 1995. Madrid: Centro de Publicaciones del Ministerio de Educación y Cultura.
- JORNET, J.M. (1987) *Una aproximación teórico-empírica a los métodos de medición de referencia criterial*. Tesis Doctoral. Valencia: Universitat de València.
- JORNET, J.M. y SUÁREZ, J.M. (1989a): «Conceptualización del Dominio educativo desde una perspectiva integradora en Evaluación Referida al Criterio». *Bordón*. 41, 2, 237-275.
- JORNET, J.M. y SUÁREZ, J.M. (1989b): «Revisión de Modelos y Métodos en la determinación de estándares y en el establecimiento de un Punto de corte en Evaluación Referida al Criterio (ERC)». *Bordón*. 41, 2, 277-301.
- JORNET, J.M. y SUÁREZ, J.M. (1994) Evaluación Referida al Criterio. Construcción de un Test Criterial de Clase. En V. GARCÍA HOZ (Dir.) *Problemas y Métodos de Investigación en Educación Personalizada*. Madrid: Rialp.
- JORNET, J.M., SUÁREZ, J.M., GONZÁLEZ SUCH, J., PÉREZ CARBONELL, A. y FERRÁNDEZ, M.R. (1993) *Evaluation Report of the Project: Communication and Presentation Skills for Technological Transfer Agents*. Euro-Innovations-Manager. Valencia: ADEIT/IMPIVA/CEEI.
- KINGSBURY, G.G. y ZARA, A.R. (1989) Procedures for selecting items for computerized adaptive testing. *Applied Measurement in Education* 2(4), 359-375.
- LADUCA, A., STAPLES, W.I., TEMPLETON, B. y HOLZMAN, G.B. (1986) Ítem modelling procedure for constructing content-equivalent multiple-choice questions. *Medical Education*, 20, 53-56.
- LINN, R.L. (1990) Has Ítem Response Theory increased the Validity of Achievement Test scores? *Applied Measurement in Education*, 3, 2, pp. 115-141.
- LIVINGSTON, S.A. y ZIEKY, M.J. (1982) *Passing Scores*. Princeton N.J.: ETS.
- MARI, R. (1996) *Evaluación del Rendimiento en Matemáticas: adaptación de la Escala Key Math-R*. Tesis Doctoral (en prensa: microficha). Valencia: Universitat de València.
- MARTÍNEZ RIZO, F. (1996) *La calidad de la educación en Aguascalientes. Diseño de un sistema de monitoreo*. Aguascalientes, México: Universidad Autónoma de Aguascalientes (UAA)-Instituto de Educación de Aguascalientes (IEA).
- NACCARATO, R.W. (1988) *A guide to item banking in Education* (3ª ed.) Portland, O.: Northwest Regional Education Laboratory.
- ORDEÑANA, B. (1991) Funcionamiento diferencial de los ítems: una aplicación al campo de las diferencias entre sexos. *Revista de Investigación Educativa*, 9, 17, 119-128.
- OSTERLIND, S.J. (1992) *Constructing test items*. (2ª ed.). Boston: Kluwer.
- PÉREZ JUSTE, R. y MARTÍNEZ ARAGÓN, L. (1989) *Evaluación de centros y calidad educativa*. Madrid: Cincel.
- POSTLETHWAITE, T.N. (1987) Introduction: Special issue on the Second IEA Study. *Comparative Educational Review*. 31(1), 150-158.

- RECKASE, M.D. (1979) Unifactor latent trait models applied to multifactor tests. *Journal of Educational Statistics*, 4, 207-230.
- RIVAS F. et al. (1986): *Proyecto Valencia: Objetivos básicos de aprendizaje en los Ciclos y Areas de Lenguaje y Matemáticas en la EGB. Una aproximación de Evaluación Referida al Criterio*. Valencia: Servicio de Estudios y Publicaciones Universitarias, S.A.
- RIVAS, F. JORNET, J.M. y SUÁREZ J.M. (1995) Evaluación del aprendizaje escolar: claves conceptuales y metodológicas básicas. En F. SILVA (De.): *Evaluación psicológica en niños y adolescentes*. Madrid: Síntesis.
- RODRÍGUEZ LAJO, M. (1986) Evaluación del rendimiento criterial vs. Normativo. Modelo de evaluación FCO. *Revista de Investigación Educativa*, 3, 6, 304-321.
- ROID, G.H. y HALADYNA, T.M. (1982) *A technology of test-item writing*. New York : Academic Press.
- SAMEJIMA, F. (1974) Normal ogive model on the continuous response level in the multidimensional latent space. *Psychometrika*, 39, 111-121.
- SHAVELSON, R.J. y WEBB, N.M. (1991) *Generalizability Theory. A Primer*. Newbury Park, CA.: SAGE.
- SPECIAL STUDY PANEL ON EDUCATION INDICATORS (SSPEI) (1991) *Education counts. An indicator system to monitor the nation's educational health*. Washington : National Center for Educational Statistics. USA Department of Education.
- TEJEDOR, F.J.; GARCÍA VALCÁRCEL, A. y RODRÍGUEZ CONDE, M.J. (1994) Perspectivas metodológicas en la evaluación de programas en el ámbito educativo. *Revista de Investigación Educativa*, 23, 93-128.
- TOBOSO J. (1995a): Fundamentos teóricos del proceso evaluador desde el marco curricular de la LOGSE. En AIDIPE (Comp.): *Estudios de Investigación Educativa en Intervención Psicopedagógica*. Valencia: AIDIPE.
- TOBOSO J. (1995b): Estudio empírico sobre la Evaluación de componentes cognitivos en la Resolución de problemas. En AIDIPE (Comp.): *Estudios de Investigación Educativa en Intervención Psicopedagógica*. Valencia: AIDIPE.
- TOCA, M.T. y TOURON, J. (1989) Factores del Rendimiento Académico en los Estudios de Arquitectura. *Revista de Investigación Educativa*, 7, 14, 31-47.
- TOURON, J. (1985) La predicción del rendimiento académico: procedimientos, resultados e implicaciones. *Revista Española de Pedagogía*, 169-170, 473-495.
- WEISS, D.J. y YOES, M.E. (1991) Ítem Response Theory. En R.K. HAMBLETON y J.N. ZAAL (eds.) *Advances in Educational and Psychological Testing*. Boston, MA: Kluwer
- WEST, A.; PENNELL, H.; THOMAS, S. y SAMMONS, P. (1995) Educational performance indicators. *EERA Bulletin*, 1, 3, 3-11.

REVISTA
INVESTIGACIÓN
EDUCATIVA

BOLETÍN SUSCRIPCIÓN

Para suscribirse llene este boletín y devuélvalo a:

REVISTA DE INVESTIGACIÓN EDUCATIVA

Dpto. «Métodos de Investigación y Diagnóstico en Educación»

Facultad de Pedagogía Paseo del Valle de Hebrón, 171, 2ª planta
08035 - BARCELONA (Spain)

Nombre

D.N.I. o N.I.F.

Dirección

Población C.P.

País Teléfono

Coste de la inscripción:

Individual: 4.000 ptas.

Institucional: 6.000 ptas.

Números sueltos: 2.000 ptas.

Indicar n.º deseado:

Números extras: 2.500 ptas.

Indicar n.º deseado:

(Fecha y Firma)

BOLETÍN DE DOMICILIACIÓN BANCARIA

Señores,

Les agradeceré que con cargo a mi cuenta/libreta atiendan los recibos que les presentará la *Revista de Investigación Educativa*, como pago de mi suscripción a la misma.

Titular de la cuenta

Banco/Caja

N.º de cuenta N.º de libreta

Agencia

Población

(Fecha y Firma)

AIDIPE

Asociación Interuniversitaria de Investigación Pedagógica, creada en 1987

FINES DE LA ASOCIACIÓN

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

DERECHOS DE LOS SOCIOS

Los socios numerarios y fundadores tendrán los siguientes derechos:

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

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

SEMINARIOS

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

PUBLICACIONES

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

Para más información, ver hoja de inscripción adjunta.

Para asociarse llene las dos partes de este boletín y devuélvalo a:

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

Dpto. «Métodos de Investigación y Diagnóstico en Educación»

Facultad de Pedagogía Paseo del Valle de Hebrón, 171, 2ª planta

08035 - BARCELONA (Spain)

Cuota de suscripción anual 5.000 ptas.

DATOS PERSONALES

Nombre y Apellidos

D.N.I. o N.I.F.

Dirección

Población C.P.....

Provincia Teléfono ().....

Deseo asociarme desde el día de de 19.....

DPTO. TRABAJO **CENTRO TRABAJO**.....

Situación profesional Dist. Universitario.....

DATOS BANCARIOS

Titular de la cuenta

Banco/Caja

Domicilio Agencia

Población C.P.....

(Firma)

CODIGO CUENTA CLIENTE			
Entidad	Oficina	D.C.	Núm. de Cuenta

Señores,

Les agradeceré que con cargo a mi cuenta/libreta atiendan los recibos que les presentará la **Asociación Interuniversitaria de Investigación Pedagógica Experimental** como pago de mi cuota de asociado.

Titular de la cuenta

Banco/Caja

Domicilio Agencia

Población C.P.....

(Fecha y Firma)

CODIGO CUENTA CLIENTE			
Entidad	Oficina	D.C.	Núm. de Cuenta

NORMAS PARA LAS COLABORACIONES

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

1. Los trabajos deberán ser originales y versar sobre investigación educativa.
2. El autor o autora deberá enviar un original y tres copias mecanografiadas a la redacción de la Revista.
3. El autor o autora deberá especificar debajo del nombre del artículo su dirección profesional y su E.mail. Si son varios autores/as, se adjuntará el E.mail del primer autor como mínimo.
4. La extensión máxima de los trabajos no deberá exceder las 25 páginas en DIN-A4, a doble espacio y numeradas. Se acompañará un abstract de 100 a 150 palabras en inglés y español.
5. Con la finalidad de simplificar el proceso de confección de la revista y, sobre todo, de reducir al máximo las erratas y costos, se ruega a los autores/as enviar el trabajo, además de las copias en papel, en disquette compatible. Se aceptan los siguientes programas de tratamiento de textos: Word Star, Word Perfect, MSWord y otros.
6. Los trabajos recibidos serán sometidos a informe del Consejo Asesor de la Revista (quien decidirá sobre su publicación o no).
7. Los trabajos recibidos para su publicación en la revista estarán sujetos a un doble proceso de filtro para garantizar la objetividad del juicio. En primer lugar, existirá un proceso de separación de las identificaciones de los autores/as de los trabajos candidatos a ser publicados junto con la asignación de códigos identificativos. Además, existirá un proceso completamente independiente de elección de jueces entre el Consejo Asesor. Ambos procesos serán totalmente independientes y transparentes entre sí, de forma que al asignar un conjunto de jueces del Consejo Asesor se desconoce cualquier información sobre el artículo, salvo el conjunto de descriptores sobre su contenido.
8. Se notificará a sus autores/as los trabajos aceptados para su publicación.
9. Para la redacción de los trabajos se recomienda que sigan las normas de la A.P.A. (American Psychology Association). (Ver adaptación publicada en el n.º 19, 1.º semestre 1992 de R.I.E.).
10. Las fichas resumen se cumplimentarán de acuerdo con el modelo propuesto e impreso por la Revista, a cuya sede podrá solicitar los ejemplares.
11. Los números se cierran los días 1 de mayo y 1 de noviembre de cada año.
12. La R.I.E. es de carácter semestral.

REDACCIÓN: REVISTA DE INVESTIGACIÓN EDUCATIVA
Dpto. de Métodos de Investigación y Diagnóstico en Educación
Facultad de Pedagogía
Paseo del Valle Hebrón, 171, 2ª planta
08035 BARCELONA (Spain)