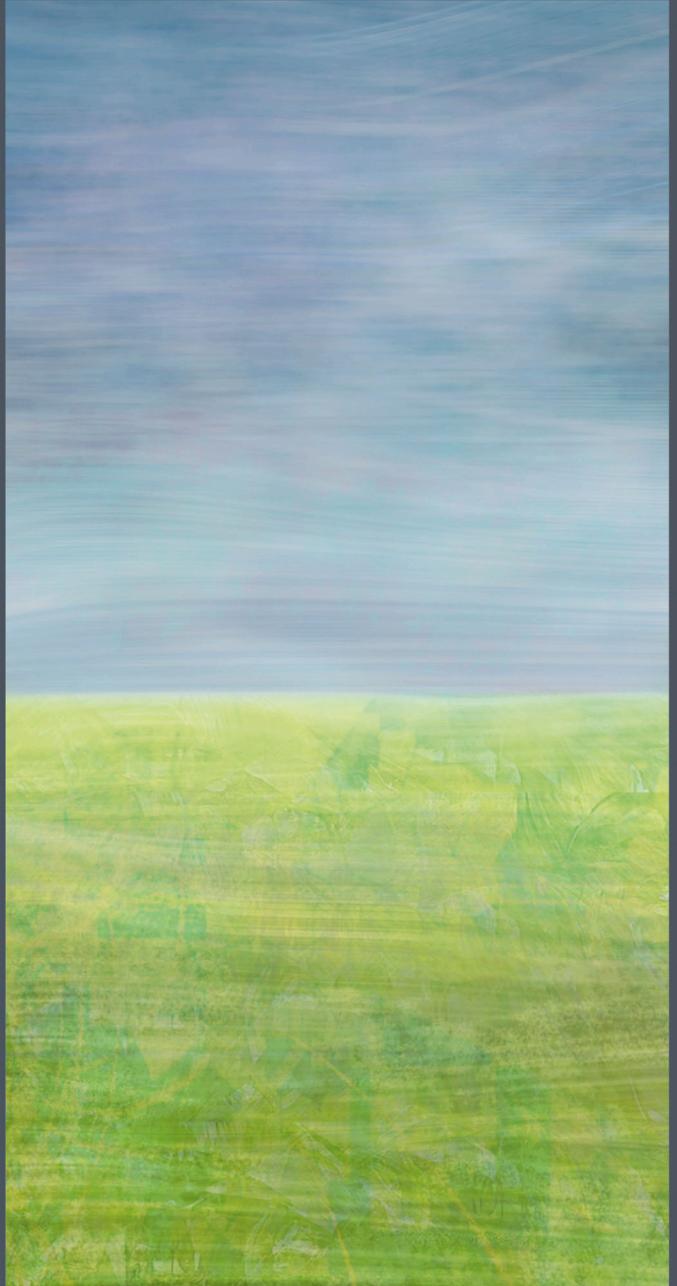


# revista de **e**DUCCIÓN

Nº 348 ENERO - ABRIL 2009

El valor añadido en educación



GOBIERNO  
DE ESPAÑA

MINISTERIO  
DE EDUCACIÓN,  
POLÍTICA SOCIAL Y DEPORTE



**revista de**  
**e**EDUCACIÓN



**Nº 348 ENERO-ABRIL 2009**

**revista de**  
**EDUCACIÓN**  
**Nº 348 enero-abril 2009**

Revista cuatrimestral

Fecha de inicio: 1952



**MINISTERIO DE EDUCACIÓN, POLÍTICA SOCIAL Y DEPORTE**  
**SECRETARÍA GENERAL DE EDUCACIÓN Y FORMACIÓN**

**Instituto de Evaluación**  
San Fernando del Jarama, 14  
28002 Madrid  
Teléfono (+34) 91 745 92 00  
Fax (+34) 91 745 92 49  
[redaccion.revista@mepsyd.es](mailto:redaccion.revista@mepsyd.es)

Edita:  
© SECRETARÍA GENERAL TÉCNICA  
Subdirección General de Información y Publicaciones

Catálogo de publicaciones del Ministerio  
<http://www.mepsyd.es/>

Suscripciones y venta: [publicaciones@mec.es](mailto:publicaciones@mec.es)

Catálogo general de publicaciones oficiales  
[www.060.es](http://www.060.es)

Edición completa en  
[www.revistaeducacion.mec.es](http://www.revistaeducacion.mec.es)

Fecha de edición: 2008  
NIPO: 651-08-003-9  
ISSN: 0034-8082  
Depósito Legal: M.57/1958

Diseño: Dinarte S.L.  
Maqueta e imprime: Estilo Estugraf Impresores, S. L.  
Pol. Ind. Los Huertecillos, nave 13 - 28350 CIEMPOZUELOS (Madrid)  
[estugraf@terra.es](mailto:estugraf@terra.es)

## CONSEJO DE DIRECCIÓN/MANAGING BOARD

### PRESIDENTE/CHAIR

Eva Almunia Badía  
Secretaría de Estado de Educación y Formación

### VOCALES/MEMBERS

Javier Vidal García  
Director General de Evaluación y Ordenación del Sistema Educativo

Rosa Peñalver Pérez  
Directora General de Cooperación Territorial

Miguel Soler Gracia  
Director General de Formación Profesional

Felipe Petriz Calvo  
Director General de Universidades

José Canal Muñoz  
Secretario General Técnico

Enrique Roca Cobo  
Director del Instituto de Evaluación

Eduardo Cobo Arango  
Director del Centro de Investigación  
y Documentación Educativa (CIDE)

## CONSEJO EDITORIAL/EDITORIAL BOARD

### DIRECTOR/CHAIR

Enrique Roca Cobo

### EDITORA JEFE/EDITOR IN CHIEF

Consuelo Vélaz de Medrano Ureta

### EDITORES ASOCIADOS/ASSOCIATED EDITORS

Juan Manuel Escudero (Universidad de Murcia);  
Mariano Fernández Enguita (Universidad de Salamanca);  
Juan Manuel Moreno Olmedilla (Senior Education Specialist. Banco Mundial);  
Carlos Marcelo (Universidad de Sevilla); Sebastián Rodríguez Espinar  
(Universidad Autónoma de Barcelona); Juan Carlos Tedesco (UNESCO);  
Alejandro Tiana Ferrer (Universidad Nacional de Educación a Distancia).

## REDACCIÓN/ASSISTANT EDITORS/MANUSCRIPTS

Jefe de Redacción: M<sup>a</sup> Jesús Pérez Zorrilla

### Equipo de Redacción

Mercedes Díaz Aranda  
Beatriz González Dorrego  
Fernando Castro Vega  
Cristina Jiménez Noblejas  
Nuria Manzano Soto (apoyo al proceso editor)  
Beatriz Pinto González (Página web)  
Gúdula Pilar García Angulo (coordinadora del diseño de la portada)

## CONSEJO ASESOR/EDITORIAL ADVISORY BOARD

### Internacional

Aaron Benavot (EFA/UNESCO. Jerusalem University); Abdeljalil Akkari (Profesor de la Universidad de Ginebra, Suiza); Mark Bray (Director, International Institute for Educational Planning. IIEP/UNESCO; Member of Advisory Board, Comparative Education Journal); José Joaquín Brunner (Universidad Diego Portales, Chile); Andy Hargreaves (Lynch School of Education. Boston College. Editor-in-Chief of the Journal of Educational Change); Seamus Hegarty (President, International Association for the Evaluation of Educational Achievement, IEA); Felipe Martínez Rizo (Director, Instituto Nacional para la Evaluación de la Educación, México); Jaap Scheerens (University of Twente, Netherlands. INES Project, OCDE); Andreas Schleicher (Head of the Indicators and Analysis Division, Directory for Education, OCDE)

### Nacional

Teresa Aguado (UNED); Margarita Bartolomé (U. de Barcelona); Jesús Beltrán Llera (U. Complutense); Antonio Bolívar (U. de Granada); Josefina Cambra (Colegios de Doctores y Licenciados); Anna Camps (U. Autónoma de Barcelona); Colectivo Ioé (Madrid); César Coll (U. de Barcelona); Agustín Dosil (U. de Santiago); Gerardo Echeita (U. Autónoma de Madrid); José Manuel Esteve (U. de Málaga); Joaquín Gairín (U. Autónoma de Barcelona); M<sup>a</sup> Angeles Galino (U. Complutense); J. L. García Garrido (UNED); José Luis Gaviria (U. Complutense); Daniel Gil (U. de Valencia); José Gimeno Sacristán (U. de Valencia); Fuensanta Hernández Pina (U. de Murcia); Carmen Labrador (U. Complutense); Ramón L. Façal (IES Pontepedriña, Santiago de Compostela. Revista Iber); Miguel López Melero (U. de Málaga); Elena Martín (U. Autónoma de Madrid); Miquel Martínez (U. de Barcelona); Rosario Martínez Arias (U. Complutense); Mario de Miguel (U. de Oviedo); Inés Miret (Neturity, Madrid); Gerardo Muñoz (Inspección de Madrid); Gema Paniagua (E. Atención Temprana, Leganés); Emilio Pedrinaci (IES El Majuelo, Sevilla); Ramón Pérez Juste (UNED); Gloria Pérez Serrano (UNED); Ignacio Pozo (U. Autónoma de Madrid); M<sup>a</sup> Dolores de Prada (Inspección); Joaquim Prats (U. de Barcelona); Manuel de Puelles (UNED); Tomás Recio (U. de Cantabria); Luis Rico (U. de Granada); Juana M<sup>a</sup> Sancho (U. de Barcelona); Consuelo Uceda (Colegio La Navata, Madrid); Mercedes Vico (U. de Málaga); Florencio Villarroya (IES Miguel Catalán, Zaragoza. Revista Suma); Antonio Viñao (U. de Murcia)



## Presentación

La REVISTA DE EDUCACIÓN es una publicación científica del Ministerio de Educación y Ciencia español. Fundada en 1940, y con el título de *Revista de Educación* desde 1952, es un testigo privilegiado de la evolución de la educación en las últimas décadas, así como un reconocido medio de difusión de los avances en la investigación y la innovación educativas, tanto desde una perspectiva nacional como internacional. Actualmente está adscrita al Instituto de Evaluación de la Secretaría General de Educación y es editada por la Subdirección General de Información y Publicaciones.

Cada año se publican tres números ordinarios y uno extraordinario dedicado a un tema de interés. Los números ordinarios agrupan las colaboraciones en cuatro secciones: *Monográfica*, *Investigaciones y estudios*, *Informes y ensayos*, y *Experiencias educativas (innovación)*, todas ellas sometidas a evaluación externa. En el primer número del año se incluyen, además, un índice bibliográfico, un editorial que recoge las principales estadísticas del proceso de editor de ese período y los índices de impacto, así como el listado de evaluadores externos.

Desde 2006 la Revista se publica en doble formato, impreso y electrónico. La edición impresa incluye los artículos de la sección monográfica en toda su extensión, los resúmenes de los artículos del resto de las secciones en español e inglés y un índice de los libros reseñados y recibidos. Por su parte, la edición electrónica incluye todos los artículos y reseñas completos, y es accesible a través de la página web ([www.revistaeducacion.mepsyd.es](http://www.revistaeducacion.mepsyd.es)), en la que además se incluye otra información de interés sobre la Revista. Los números extraordinarios se publican íntegros en ambos formatos. En todos los números, la versión completa se encuentra también en el CD que acompaña a la Revista impresa.

*Revista de Educación* evalúa, selecciona y publica trabajos enmarcados en líneas de investigación consolidadas, principalmente sobre: *metodologías de investigación y evaluación educativa; análisis de sistemas educativos y políticas públicas; evolución e historia de los sistemas educativos contemporáneos; reformas e innovaciones educativas; calidad y equidad en educación; currículo; didáctica; organización y dirección escolar; atención a la diversidad*

*y educación inclusiva; orientación educativa y tutoría; selección, formación y desarrollo profesional del profesorado; cooperación internacional para el desarrollo de la educación.*

La *Revista de Educación* aparece en los siguientes medios de documentación bibliográfica:

- *Bases de datos españolas*: BEG (GENCAT), ISOC, PSEDISOC, PSICODOC, DIALNET y CEDUS (Centro de Documentación Universitaria), RESH (Revistas Españolas de Ciencias Sociales y Humanas), REDINED (Red de Bases de Datos de Información Educativa).
- *Bases de datos internacionales*: Social Sciences Citation Index® (SSCI), Social Scisearch®, Journal Citation Reports/Social Sciences Edition, European Reference Index for the Humanities (ERIH), Ulrich's Periodicals Directory, LATINDEX (Iberoamericana); PIO (Periodical Index Online, Reino Unido), IPSA (International Political Science Abstracts), IRESIE (México); ICIST (Canadá); HEDBIB (UNESCO-Higher Education Bibliography); SWETSNET (Holanda).
- *Catálogos nacionales*: Consejo Superior de Investigaciones Científicas (CSIC-ISOC); Red de Bibliotecas Universitarias (REBIUN); Centro de Investigación y Documentación Educativa (CIDE); Catálogo Colectivo de Publicaciones Periódicas en Bibliotecas Españolas (Ministerio de Educación y Ciencia).
- *Catálogos internacionales*: Online Computer Library Center (USA); Library of Congress (LC); The British Library Current Serials Received; King's College London; University of London-ULRLS; Colectif National Français; Centro de Recursos Documentales e Informáticos de la Organización de Estados Iberoamericanos (OEI).

**La Revista no comparte necesariamente las opiniones  
y juicios expuestos en los trabajos firmados.**

## Presentation

REVISTA DE EDUCACIÓN is a scientific journal published by the Spanish Ministry of Education and Science. Founded in 1940, and since 1952 called *Revista de Educación*, it has been a privileged witness of the development of education in the last decades, and an acknowledged means for the dissemination of education research and innovation, both from a national and international perspectives. It is currently assigned to the Institute of Evaluation within the General Secretary of Education and it is published by the General Directorate of Information and Publications of the Ministry of Education and Science.

Four issues are published each year, three regular ones and one focused on an especial topic. Regular issues include four sections: Monograph, Research and Studies, Reports and Essays and Education Experiences (innovation), all of them submitted to referees. In the first issue of the year there is also an index of bibliography, a report with statistic information about the journal process of this period and the impact factors, as well as a list of our external advisors.

Since 2006, *Revista de Educación* is published in a double format, paper and electronic. The paper edition includes all the articles in the especial section, the abstracts of articles pertaining to the rest of sections, and an index of reviewed and received books. The electronic edition contains all articles and reviews of each issue, and it is available through this web page ([www.revistaeducacion.mec.es](http://www.revistaeducacion.mec.es)), where it is possible to find more interesting information about the journal. The focused-topic issues are published full-length in both formats. In all the issues, a full-length version of regular issues is also available in the CD that comes together with the paper edition.

*Revista de Educación* assesses, selects and publishes studies framed in wellestablished lines of research, mainly: *methodologies of education investigation and assessment; analysis of education systems and public policies; evolution and history of contemporary education systems; education reforms and innovations; quality and equity in education; curriculum; didactics; school organization and management; attention to diversity*

*and inclusive education; educational guidance and tutorship; teacher selection, training and professional development; internacional cooperation for the development of education.*

*Revista de Educación* is available through the following data bases:

- *Spanish databases:* BEG (GENCAT), ISOC, PSEDISOC, PSICODOC, DIALNET y CEDUS (Centro de Documentación Universitaria), RESH (Revistas Españolas de Ciencias Sociales y Humanas), REDINED (Red de Bases de Datos de Información Educativa).
- *International databases:* Social Sciences Citation Index® (SSCI), Social Scisearch®, Journal Citation Reports/Social Sciences Edition, European Reference Index for the Humanities (ERIH), Ulrich's Periodicals Directory, LATINDEX (Iberoamericana); PIO (Periodical Index Online, UK), IPSA (International Political Science Abstracts), IRESIE (Mexico); ICIST (Canada); HED-BIB (UNESCO-Higher Education Bibliography); SWETSNET (Netherlands).
- *National Catalogues:* Consejo Superior de Investigaciones Científicas (CSIC); Red de Bibliotecas Universitarias (REBIUN); Centro de Investigación y Documentación Educativa (CIDE); Catálogo Colectivo de Publicaciones Periódicas en Bibliotecas Españolas (Ministerio de Educación y Ciencia).
- *International Catalogues:* Online Computer Library Center (USA); Library of Congress (LC); The British Library Current Serials Received; King's College London; University of London-ULRLS; Colectif National Français; Centro de Recursos Documentales e Informáticos de la Organización de Estados Iberoamericanos (OEI).

***Revista de Educación* does not necessarily agree with opinions and judgments maintained by authors.**

## Monográfico

### El valor añadido en educación Value-added in education

ROSARIO MARTÍNEZ ARIAS, JOSÉ LUIS GAVIRIA SOTO Y MARÍA CASTRO MORERA. Concepto y evolución de los modelos de valor añadido en educación.....	15
ANDREW RAY, HELEN EVANS Y TANYA McCORMACK. El uso de modelos nacionales de valor añadido para la mejora de las escuelas británicas .....	47
YEOW MENG THUM. No Child Left Behind: retos metodológicos y recomendaciones para la medida del progreso anual adecuado .....	67
EVA BAKER. Consideraciones de validez prioritaria para la evaluación formativa y de rendición de cuentas .....	91
MARÍA CASTRO MORERA, COVADONGA RUIZ Y ESTHER LÓPEZ. Forma básica del crecimiento en los modelos de valor añadido: vías para la supresión del efecto de regresión .....	111
MARÍA EUGÉNIA FERRÃO. Sensibilidad de las especificaciones de los modelos de valor añadido: midiendo el estatus socioeconómico.....	137
JOSÉ LUIS GAVIRIA, M <sup>a</sup> CHANTAL BIENCINTO, ENRIQUE NAVARRO. Invarianza de la estructura de covarianzas de las medidas de rendimiento académico en estudios longitudinales en la transición de la Educación Primaria a Secundaria .....	153
LUIS LIZASOAIN HERNÁNDEZ Y LUIS JOARISTI OLARIAGA. Análisis de la dimensionalidad en los modelos de valor añadido: estudio de las pruebas de matemáticas empleando técnicas factoriales y métodos no paramétricos basados en TRI (Teoría de Respuesta al Ítem).....	175
ÁNGELES BLANCO, CORAL GONZÁLEZ Y XAVIER ORDOÑEZ. Patrones de correlación entre medidas de rendimiento escolar en evaluaciones longitudinales: un estudio de simulación desde un enfoque multinivel .....	195
ROSARIO MARTÍNEZ ARIAS. Usos, aplicaciones y problemas de los modelos de valor añadido en educación .....	217

## Investigaciones y Estudios

ÓSCAR MEDINA FERNÁNDEZ Y FLORENTINO SANZ FERNÁNDEZ. Los sistemas de reconocimiento y acreditación de los aprendizajes no formales e informales: referencias históricas, funciones socioeducativas y perspectiva teórica.....	253
GUILLERMO RUIZ. La nueva reforma educativa argentina según sus bases legales .....	283
ANTONIO GARCÍA GUZMÁN. Los procesos cognitivos y metacognitivos en la composición escrita del alumnado de etnia gitana con desventaja socioeducativa .....	309
JULIÁN LÓPEZ YÁÑEZ Y MARITA SÁNCHEZ MORENO. Mujeres agentes de cambio en la dirección de organizaciones universitarias .....	331
COVADONGA RUIZ DE MIGUEL. Las escuelas eficaces: un estudio multinivel de factores explicativos del rendimiento escolar en el área de matemáticas.....	355
TERESA MAURI MAJÓS, ROSA COLOMINA ÁLVAREZ E INÉS DE GISPERT. Diseño de propuestas docentes en TIC en la Enseñanza Superior: nuevos retos y principios de calidad desde una perspectiva socioconstructivista .....	377
FRANCISCO CÓRDOBA ALCAIDE, ROSARIO ORTEGA RUIZ Y ALFONSO PONTES PEDRAJAS. Los universitarios de ciencias ante la docencia en Secundaria como experiencia profesional .....	401
JUAN ANTONIO MORENO MURCIA, JOSÉ ANTONIO VERA LACÁRCEL Y EDUARDO CERVELLO GIMENO. Efectos de la cesión de responsabilidad de la evaluación en la motivación y la competencia percibida en el aula de educación física .....	423

## Ensayos e informes

RAFAEL PÉREZ GALÁN. Los nuevos retos del aprendizaje social en niños con necesidades educativas especiales. El aprendizaje en comunidad o la comunidad de aprendizaje en el aula .....	443
ENCARNACIÓN SÁNCHEZ LISSEN. Mitos y realidades en la carrera docente.....	465
ANA BENITO MARTÍN. La pedagogía no tiene la culpa: un análisis de los problemas de la educación en España .....	489

## Experiencias educativas (innovación)

MANUEL CUADRADO, MARÍA EUGENIA RUIZ Y MERCEDES COCA. Participación y rendimiento del estudiante universitario en un proyecto docente interdisciplinar, bilingüe y virtual .....	505
---	-----

## Índice bibliográfico

Índice bibliográfico de la <i>Revista de Educación</i> . Año 2008 (números 345-346-347 y extraordinario) .....	521
--	-----

Recensiones y libros recibidos.....	543
-------------------------------------	-----



**Monográfico**

**El valor añadido en educación**

**Value-added in education**



# Concepto y evolución de los modelos de valor añadido en educación

## Concept and evolution of educational value-added models

Rosario Martínez Arias

*Universidad Complutense de Madrid. Facultad de Psicología. Departamento de Metodología de las Ciencias del Comportamiento. Madrid, España*

José Luis Gaviria Soto

María Castro Morera

*Universidad Complutense de Madrid. Facultad de Educación. Departamento de Métodos de Investigación y Diagnóstico en Educación (MIDE). Madrid, España*

### Resumen

Los modelos de valor añadido son un conjunto de técnicas estadísticas complejas que utilizan datos de puntuaciones de tests de los estudiantes de varios años, para estimar los efectos de las escuelas individuales. Los modelos intentan aislar la contribución de la escuela al desarrollo del aprendizaje de los alumnos. Existen diversas variaciones de los modelos que se utilizan en trabajos de investigación y en evaluaciones prácticas. En este artículo se presenta el concepto de valor añadido de la escuela y su historia y evolución. Se establece el origen y desarrollo de los modelos en torno a tres líneas principales: la investigación sobre la efectividad de las escuelas, las críticas derivadas de los informes actuales de rendición de cuentas y el desarrollo de los modelos estadísticos multinivel. Se explica la motivación de su uso para evitar algunos problemas frecuentes en las presentaciones de los resultados de las escuelas y en su ordenación, tal como sucede con las «Tablas de Liga en Inglaterra» y los informes de «Progreso Anual Adecuado» consecuencia de la aplicación de la *No Child Left Behind* en USA. Los principales problemas analizados son los relacionados con la metodología transversal de las evaluaciones

y los sesgos de selección derivados de las características de los estudiantes en el ingreso, variables sociodemográficas y factores contextuales de las escuelas. En resumen, se exponen las principales aproximaciones existentes para medir los efectos de las escuelas y se concluye que la aplicación de los procedimientos de Valor Añadido representa una importante promesa para la evaluación de las escuelas.

*Palabras clave:* investigación sobre efectividad de las escuelas, rendición de cuentas, evaluación de escuelas, modelos de valor añadido.

### **Abstract**

Value-added modelling is a collection of complex statistical techniques which use multiple years of students' test score data to estimate the effects of individual schools. The models attempt to isolate the contributions of schools to student learning development. Several variations of these models are being applied in the research literature and in practical school assessments.

In this paper we introduce the concept of value-added and its history and evolution. We establish the origin and the development of the models around three main lines: the school effectiveness research, the criticism on actual school accountability reports, and the development of multilevel statistical models. We explain the motivation to use these models in order to prevent some problems with the reporting of school results and school ratings, such as the League Tables in England and the Adequate Yearly Progress from the *No Child Left Behind* in USA. The main analyzed problems are those related to the cross-sectional methodology of most of assessments, and the selection biases derived from the student characteristics at the intake, socio-demographic variables and contextual factors of schools. We summarize the principal existing modeling approaches for measuring school effects and we conclude that the application of value-added procedures holds considerable promise for school assessment.

*Key Words:* school effectiveness research, school accountability, school assessment, value-added models.

## **Introducción**

Muchos gobiernos están en la actualidad insatisfechos con los niveles de rendimiento de sus estudiantes y reciben presiones crecientes para mejorar la eficacia y eficiencia del sistema educativo. Esta preocupación está abonada por la investigación aplicada de la economía de la educación que establece que la prosperidad económica y social

de los individuos y de las naciones radica en la educación y formación (Hanushek, 2005; Hanushek y Wossman, 2006). El nivel de logros educativos de un país suele considerarse como un indicador de sus reservas de «capital humano» o disponibilidad de potenciales trabajadores con educación y destrezas suficientes. Además, la sociedad debe asegurar la igualdad de acceso a las oportunidades educativas para lograr niveles satisfactorios de bienestar individual y social. Es decir, la educación proporcionada debe ser equitativa, para lo que deben reducirse las brechas de rendimiento entre diferentes grupos de estudiantes. Esta preocupación por alcanzar adecuados niveles de rendimiento se ha extendido a la sociedad en general, como pone de relieve el creciente interés mediático por los resultados de las encuestas internacionales.

Tanto los partidarios de la rendición de cuentas de las escuelas como los no partidarios reconocen que un ingrediente fundamental para la mejora de las escuelas y del sistema educativo en general es un buen sistema de evaluación de las escuelas, que permita disponer de datos en los momentos adecuados. Estos datos servirán para emprender acciones de mejora de las escuelas y para la planificación de las reformas educativas.

Es en este contexto en el que la metodología conocida como modelos de Valor Añadido (en adelante VA) se ha recibido con entusiasmo y con grandes expectativas por parte de los responsables de la educación.

Aunque no existe una definición única de lo que es el VA de una escuela, suele considerarse como su contribución al progreso neto de los estudiantes hacia objetivos de aprendizaje establecidos, una vez eliminada la influencia de otros factores ajenos a la escuela que pueden contribuir a dicho progreso. Los modelos de VA son un conjunto de procedimientos estadísticos que se utilizan para hacer inferencias sobre la eficacia de las escuelas y de los profesores que ponen el acento en las ganancias de los estudiantes en el tiempo. Tienen en común el seguimiento de la trayectoria de los estudiantes analizando las medidas de los resultados de dos o más años. Estos datos de la evolución de los estudiantes se transforman en indicadores de la eficacia de la escuela o del profesor. La idea es simple: las escuelas más eficaces son aquellas cuyos estudiantes ganan más, mientras que las menos eficaces son las que ganan menos. Algunos de estos modelos, los denominados contextualizados, ajustan los resultados por variables socioeconómicas y demográficas con objeto de hacer más justas las comparaciones entre escuelas.

La importancia atribuida a los modelos de VA en la actualidad se debe a un gran interés por hacer a las escuelas responsables de sus rendimientos rindiendo cuentas a la sociedad de los aprendizajes de sus estudiantes. Este interés se encuentra fomentado por los procesos de descentralización y de autonomía creciente para las escuelas que

se han emprendido en muchos países. La autonomía debe estar acompañada de evaluaciones estructuradas y sistemáticas de los resultados que proporcionen feedback a las escuelas sobre sus fuerzas y debilidades y sobre qué aspectos deben mejorar. Es en este contexto en el que los modelos de VA pueden ser de extraordinaria utilidad.

Aunque el interés por los modelos de VA para la evaluación es reciente, en realidad no son tan nuevos, sino que son el producto de diferentes líneas de investigación que marcan su desarrollo y algunos de los debates sobre su utilización. En este artículo se presentan las líneas de investigación que han llevado a los modelos de VA actuales, las razones o motivación para su implantación y los principales modelos propuestos. Limitaciones de espacio impiden entrar en una explicación detallada de los modelos. Los lectores interesados pueden encontrar información complementaria en algunas fuentes recientes como la revisión realizada por McCaffrey, Lockwood, Koretz y Hamilton (2003), el número monográfico de la revista *Journal of Educational and Behavioral Statistics* (2004, (29), 1), editado por Wainer y en las compilaciones de los symposia recientes sobre el tema de la Universidad de Maryland editadas por Lissitz (Lissitz, 2005, 2006).

## El caldo de cultivo para los modelos de VA

En este apartado y sin ánimo de exhaustividad, se presentan los antecedentes que están en la base de los modelos de VA. Se agrupan los antecedentes en tres líneas: la investigación sobre las escuelas efectivas, la política de rendición de cuentas y los avances en las técnicas estadísticas que hicieron posible la aplicación de los modelos. En realidad, es difícil separarlas, ya que están bastante unidas y son coexistentes en el tiempo. La investigación sobre la efectividad de las escuelas está más enraizada en la búsqueda de factores que determinan la calidad y la equidad de los aprendizajes; la línea que denominamos de rendición de cuentas basada en las puntuaciones de los tests puede considerarse más vinculada a los esfuerzos de evaluación de las escuelas, con consecuencia o no, dentro de las políticas de publicidad de resultados del sector público. Los desarrollos de las técnicas estadísticas están en la base de las dos líneas. En el desarrollo de los modelos también han influido la tendencia creciente a las evaluaciones a gran escala mediante tests estandarizados, nacionales e internacionales y la creciente disponibilidad de ordenadores para la creación y almacenamiento de bases de datos y de un software adecuado para su tratamiento.

## La investigación sobre la efectividad de las escuelas IEE (School Effectiveness Research)

Sus orígenes se encuentran principalmente en el Reino Unido y Estados Unidos, donde influyentes estudios realizados durante los años sesenta y setenta pusieron de relieve la escasa influencia de la escuela sobre los resultados educativos, en comparación con otros aspectos de los estudiantes, tales como las aptitudes, la etnia/raza y el estatus socioeconómico (Coleman et al., 1966; Jencks et al., 1972).

Estos tempranos estudios adolecían de un buen número de limitaciones metodológicas y la investigación posterior intentó poner de relieve la existencia de efectos escolares significativos, aún reconociendo la gran influencia del contexto socioeconómico y cultural de los estudiantes (Edmonds, 1979; Mortimore, Sammons, Stoll, Lewis y Ecob, 1988). Se intenta dar respuesta a las siguientes preguntas: *¿Pueden las escuelas ser efectivas?*, *¿Qué determina la efectividad de la escuela?* y *¿Hasta qué punto la escuela es efectiva en la reducción de las desigualdades en rendimiento debidas al origen social o étnico de los sujetos?*

Desde sus comienzos, la IEE se caracteriza por seguir una metodología fundamentalmente cuantitativa, basada en las puntuaciones de los estudiantes en tests estandarizados y por el uso de procedimientos estadísticos complejos que permitan desenmarañar de los resultados de los estudiantes la influencia de otros efectos ajenos a la escuela y a sus prácticas educativas. En un primer momento se caracterizó por el estudio de lo que se podría denominar el «rendimiento contextualizado» mediante estudios transversales realizados en un momento en el tiempo. Un ejemplo de esta metodología son los estudios PISA, en los que se intenta eliminar de los resultados otros aspectos relacionados con características socioculturales de los estudiantes. Un ejemplo de este enfoque con los datos españoles de PISA 2003 puede encontrarse en Marchesi y Martínez Arias (2006). Esta aproximación por su carácter transversal resulta insuficiente (Willms y Raudenbush, 1989; Goldstein, 1987). La investigación más reciente se caracteriza por estudiar muestras amplias de escuelas y poner el acento en la evaluación del progreso sobre el tiempo, más que en las instantáneas tomadas en un único momento derivadas de los estudios transversales. En la actualidad y siguiendo a Sammons (2006) puede considerarse la IEE como una línea de investigación que intenta separar las características de los estudiantes de los efectos de las escuelas.

No podemos entrar aquí en los principales resultados derivados de la IEE y remitimos a los lectores a Scheerens y Bosker (1997) y Teddlie y Reynolds (2000) y a la revisión de los diferentes meta-análisis realizados por Scheerens (2005). En general,

puede decirse que sus resultados son algo más alentadores que los de los tempranos estudios de Coleman y Jenks. Puede concluirse que *las escuelas establecen diferencias* (MacBeath y Mortimore, 2001; Reynolds y Creemers, 1990; Sammons, Hillman y Mortimore, 1995; Sammons y Reynolds, 1997; Scheerens, 2005; Scheerens y Bosker, 1997; Teddlie y Reynolds, 2000). Hay un porcentaje de la varianza entre escuelas (entre el 5-35%, según los estudios) que se explica por políticas y prácticas educativas y por el ambiente y clima del aprendizaje de la escuela. También se ha encontrado una escasa influencia de las variables de *inputs* o recursos en los países desarrollados, excepto en los aspectos relacionados con la formación y experiencia del profesorado, resultado coincidente con los derivados de la línea conocida como «función de producción» (Hanushek, 2003) en la Economía de la Educación. Estos datos han desplazado el acento de los *inputs* a los procesos y a los resultados en la evaluación educativa. Por lo que se refiere a las variables de procesos de las escuelas, los resultados son algo ambiguos (Muijs y Reynolds, 2000; Stevens, 2005; Zvoch y Stevens, 2003). Otro resultado bastante claro es que las escuelas efectivas responden mejor a la meta de la equidad, ya que los sujetos en desventaja progresan más en estas escuelas.

En lo que sí existe un amplio acuerdo es en que las comparaciones entre escuelas no pueden establecerse sobre los resultados brutos, sino que deben basarse en ajustes del rendimiento inicial y de otros factores relevantes y en el progreso de sus estudiantes (Goldstein et al., 1993; Goldstein y Thomas, 1996; Gray, Jesson, Goldstein, Hedger y Rasbash, 1996; Mortimore, Sammons y Thomas, 1994; Sammons, 1996). Se considera una escuela efectiva aquella en la que los estudiantes progresan más allá de lo que puede esperarse, añadiendo valor extra a los resultados de sus alumnos en comparación con otras escuelas que sirven a poblaciones que son similares en el ingreso.

## **La evaluación mediante tests al servicio de las reformas educativas y la política de rendición de cuentas**

La evaluación externa mediante tests estandarizados, ha sido constante en los diversos intentos de reformas educativas en los Estados Unidos desde los años sesenta (Hamilton, 2003; Wang, Beckett y Brown, 2006). Destacan, entre otros, los usos de los tests para el diagnóstico y monitorización del sistema educativo, representados

por el *National Assessment of Educational Progress* (NAEP), que comenzó en 1969 y que introdujo una rigurosa metodología psicométrica y estadística que tuvo una gran repercusión en las evaluaciones posteriores nacionales e internacionales. Además, las alarmas desencadenadas tras la participación de los Estados Unidos en las tempranas evaluaciones internacionales del rendimiento educativo bajo los auspicios de la *International Association for the Evaluation of Educational Achievement* (IEA) influyeron considerablemente en los posteriores desarrollos de la evaluación. Un hito importante fue la publicación en 1983 del influyente texto *A nation at risk* (National Commission of Excellence in Education, 1983). En él se insiste en los riesgos derivados de la pérdida de competitividad futura frente a otras naciones, como consecuencia de la baja calidad de la educación. Las reflexiones a que dio lugar pueden considerarse el desencadenante de diversas reformas educativas para las que resultó esencial la evaluación mediante tests, tales como el cambio de las competencias mínimas a altos estándares de rendimiento y la necesidad de disponer de datos para la monitorización y reforma del sistema. Como consecuencia, en muchos estados se iniciaron evaluaciones a gran escala, para disponer de datos para la rendición cuentas y para la mejora de las escuelas. Todo ello culminó en la legislación federal con *No Child Left Behind* (NCLB, aprobada en enero de 2002), con la exigencia de la rendición de cuentas de las escuelas en todos los estados para poder acogerse a beneficios federales.

Es importante destacar también la introducción a partir de los años noventa de la evaluación basada en *estándares de rendimiento* dentro de las reformas encaminadas a la mejora de las escuelas, sistema que recoge la NCLB. La idea fundamental de esta forma de evaluación es que los gobiernos especifican lo que los estudiantes deben saber y ser capaces de hacer en los distintos cursos y materias y estas especificaciones se recogen en los denominados *estándares de contenido* (O'Day y Smith, 1993). Los *estándares de rendimiento* representan una forma de evaluación referida a criterios que permiten establecer diferentes niveles de logro que se establecen a partir de las puntuaciones de los tests, juicios de expertos y consideraciones de política educativa. Para más información véanse Cizek y Bunch (2007) y en castellano Martínez Arias, Hernández Lloreda y Hernández Lloreda (2006). En general, esta forma de evaluación con altos estándares y la clasificación por niveles es muy bien aceptada por los profesores y por la opinión pública (Wang et al., 2006).

El uso de los tests está tan generalizado que se habla de la rendición de cuentas basada en tests, aunque evaluación y rendición de cuentas son dos procesos diferentes (Hill, Scout, DePascale, Duna y Simpson, 2006). Un aspecto importante de

estas políticas es evaluar si los estudiantes realizan progresos satisfactorios y si alcanzan los estándares de rendimiento establecidos por las autoridades. También, en aras de la equidad, se evalúa el rendimiento de distintos grupos de estudiantes, caracterizados por diferentes variables sociodemográficas (género, etnia, estatus socioeconómico).

La política de rendición de cuentas no llegó con la NCLB como algo nuevo, sino que en los últimos veinte años ya hubo un creciente interés dentro de las políticas generales de los gobiernos y estados en Estados Unidos y en otros países en la evaluación externa de las escuelas como un instrumento de supervisión útil para la mejora continua del sistema educativo (Braun y Kanjee, 2006; Kane y Staiger, 2002; Goldstein y Spiegelhalter, 1996; Hanushek y Raymond, 2004; Taylor y Nguyen, 2006). Aunque hay diferentes formas de rendición de cuentas que van desde la simple publicación de los resultados a incentivos más directos como la aplicación de recompensas y sanciones, en todos los casos, siempre es un instrumento de los políticos para ver el grado de cumplimiento de las metas, nunca una meta en sí misma. Se supone que aporta una información que puede apoyar la mejora continua de las escuelas.

Dentro de esta línea de rendición de cuentas para la mejora de los rendimientos de todos los estudiantes y de los estándares de rendimiento, la NCLB establece que todos deben alcanzar el nivel de competente en 2014, estableciendo metas anuales de progreso. También se establecen metas que reduzcan las brechas de rendimiento entre diferentes subgrupos. Impone la evaluación mediante tests anuales de Lengua y Matemáticas en todas las escuelas en los cursos de 3° a 8°. Establece consecuencias en términos de recompensas y sanciones para las escuelas según el cumplimiento o no de los objetivos. Son precisamente las consecuencias la parte más controvertida de la NCLB (Cronin, Kingsbury, McCall y Bowe, 2005; Carlson, Martínez, O'Day, Stecher, Taylor y Cook, 2007).

Los principales modelos alternativos a los de VA y que han sido objeto de numerosas críticas son el modelo de estatus y el modelo de mejora. El modelo de estatus (Goldschmidt et al., 2005; Raudenbush, 2004a) compara los resultados anuales de la escuela con el objetivo establecido (el nivel de competente), conocido como *Adequate Yearly Progress* (AYP). El modelo de mejora compara los porcentajes de estudiantes de una cohorte que alcanzan el objetivo en un curso y año particular con los porcentajes de la cohorte siguiente en el mismo curso que tiene el problema de que las diferencias en rendimiento pueden deberse a errores de muestreo y cambios en las cohortes (Hill y DePascale, 2003; Linn y Haug, 2002; Linn, Baker y Betebenner, 2002; Rouse, 2005; Zvock y Stevens, 2006).

Muchas de las críticas a los modelos de estatus y de mejora, en sus diversas variantes, se deben precisamente al uso de datos transversales (Linn, 2005; Linn y Haug, 2002; Raudenbush, 2004,b; Rouse, 2005; Stevens, 2005; Thum, 2002; Zvoch y Stevens, 2006), dada su menor susceptibilidad a las características de los estudiantes y otros factores contextuales. Hay una gran cantidad de incertidumbre asociada con las ganancias de cohortes sucesivas debida a errores de muestreo y diferencias de movilidad. El seguimiento de los estudiantes es el que permite separar los efectos sistemáticos de las escuelas de las características de los estudiantes.

Las ganancias promedio en rendimiento basadas en el seguimiento longitudinal individual de los estudiantes son las únicas que pueden proporcionar bases para eliminar (al menos en parte) las explicaciones competidoras de las diferencias en rendimiento entre escuelas (Sanders y Horn, 1994, 1998; Ballou, Sanders, y Wright, 2004; McCaffrey et al, 2003; McCaffrey, Lockwood, Koretz, Louis y Hamilton, 2004).

## La rendición de cuentas en Inglaterra

La publicación de la normativa *Every Child Matters* (1988) supuso el comienzo de la publicación de los resultados de las escuelas en tests externos realizados en determinados momentos o estadios clave. En este caso el objetivo de la evaluación es el de la identificación de escuelas efectivas e inefectivas, de cara a la introducción de mejoras y proporcionar información a los padres para la elección de centro. Esta práctica fue continuada por los gobiernos laboristas. Inicialmente los resultados se reportaban en forma de las denominadas *Tablas de Liga*, cuyo uso fue muy controvertido desde sus inicios. Los principales argumentos de sus detractores se basaban en la injusticia de los resultados brutos debido al extraordinario impacto que en ellos tenía el nivel anterior de los estudiantes y variables sociodemográficas (Goldstein y Spiegelhalter, 1996; Nuttall, Goldstein, Prosser y Rasbash, 1989; Saunders, 1999; Yang, Goldstein, Rath y Hill, 1999). Mostraron que después del ajuste de los niveles previos de los estudiantes cambiaba la ordenación de muchas escuelas y que cuando se tomaban los intervalos de confianza en torno a las medias, la mayor parte de las escuelas no se diferenciaban del promedio (Fitz-Gibbon, 1997). Estas consideraciones llevaron a una aproximación a la evaluación basada en modelos de VA para ajustar los efectos del rendimiento anterior y posteriormente al valor añadido contextualizado en 2005, incluyendo además los intervalos de confianza (Ray, Evans y McCormack, 2008, en este volumen; Schagen, 2006). Una revisión de la evolución del VA en Inglaterra puede verse en Saunders (1999).

## Los procedimientos estadísticos

El camino hacia los modelos de VA en la evaluación no habría sido posible sin el desarrollo de modelos estadísticos que permiten descomponer la variación de los resultados de los estudiantes en diferentes fuentes de variación, analizar variables procedentes de distintos niveles y tener en cuenta las dependencias de los datos de las escuelas. La mayor parte de las técnicas estadísticas paramétricas utilizadas tradicionalmente en la evaluación de los efectos de las escuelas suponen la independencia de los errores aleatorios, provocando importantes sesgos cuando no se cumple este supuesto. Por otra parte, los datos educativos tienen una estructura multinivel en la que las escuelas están anidadas en contextos, las clases en las escuelas y los estudiantes en clases y profesores. Esta estructura de los datos provoca dependencias entre las unidades de análisis, que viola el supuesto de independencia. Por otra parte, en la investigación de la eficacia de las escuelas se da el problema de integrar datos procedentes de diferentes niveles en un modelo único, integrando en la predicción del resultado predictores de diferentes niveles (De Leeuw y Meijer, 2008). Los iniciales trabajos de Coleman et al. (1966) y de Jencks et al. (1972) no tuvieron en cuenta ni la estructura de los datos, ni las dependencias. Ya en 1939 Thondike advirtió de los peligros de usar estimaciones derivadas de la correlación y regresión registradas en el nivel de grupo para hacer inferencias relativas a individuos y subgrupos, fenómeno denominado por Robinson (1950) como «falacia ecológica». Se llegó a una crítica metodológica explícita sobre las aproximaciones de nivel agregado (Burnstein, 1980), llegando en los años ochenta a las aproximaciones hoy conocidas como «modelos lineales mixtos», «modelos multinivel» o «modelos lineales jerárquicos», que ya se estaban utilizando en otras áreas (Aitkin y Longford, 1986; DeLeeuw y Kreft, 1986; Goldstein, 1987; Raudenbush y Bryk, 1986). Estos modelos permitieron la partición de la varianza en diversos niveles y la inclusión de variables predictoras que permiten explicar estas varianzas, con la inclusión de efectos de interacción. La metodología propuesta permite evitar numerosos errores de inferencia estadística (errores de tipo I, sesgos de agregación, heterogeneidad de la regresión, etc.) y separar los efectos de la escuela de otros factores de los estudiantes, llevando a los modelos de rendimiento contextualizados. Estos modelos representaron un gran paso, pero no acabaron con el debate, debido a la inestabilidad temporal de los resultados; escuelas ejemplares un año pueden ser mediocres en otro, a pesar de la supuesta estabilidad de las políticas y prácticas de las escuelas). La extensión de estos modelos a datos longitudinales (Wilms y Raudenbush, 1989) supuso

un importante avance para examinar trayectorias de desarrollo y los efectos de las escuelas sobre estas trayectorias.

En la actualidad existe un importante cuerpo de literatura en forma de libros que permite conocer estos modelos (De Leeuw y Meijer, 2008; Goldstein, 2003; Raudenbush y Bryk, 2002; Gaviria y Castro, 2005; Hox, 2002; Gelman y Hill, 2007). El uso de estas técnicas y su aplicación a la evaluación de los efectos de las escuelas no habría sido posible sin los desarrollos de software estadístico que permitan implementar los complejos procesos de estimación. Se dispone en la actualidad de numeroso software en programas específicos como MLWIN, HLM o AML, y en paquetes de programas de uso general, tanto comerciales (SAS, SPSS, SYSTAT, MPLUS, STATA-GLLAMM, SYSTAT, SPLUS), como de libre acceso (WinBUGS, R), por mencionar solamente algunos<sup>1</sup>.

## Motivación para usar los modelos de VA

Las numerosas críticas derivadas del uso de los modelos utilizados para cumplir con los requisitos de la NCLB, así como otras críticas, entre las que destacan las realizadas a las tablas de liga, condujeron a un considerable entusiasmo por los modelos de VA como una alternativa (Carey, 2004). Aunque se siguen presentando los resultados como se ha descrito en el apartado anterior, el Gobierno de Estados Unidos se hizo eco de las principales críticas y a finales de 2005 invitó a los estados a hacer propuestas de modelos basados en ganancias de los estudiantes, aprobando algunos estudios piloto que tienen en cuenta el valor añadido.

Las críticas vienen de muchos frentes, aunque la mayor parte pueden agruparse en dos grandes bloques:

- Los sesgos de selección en las escuelas.
- El uso de estudios transversales para la evaluación de las mejoras de aprendizaje.

---

<sup>1</sup> Las siglas se refieren a los nombres de los programas. Una revisión crítica de la mayor parte de los mencionados con ejemplos de aplicación puede encontrarse en el *Centre for Multilevel Modelling* de la Universidad de Bristol: <http://www.cmm.bristol.ac.uk>

## Los efectos de las escuelas y los sesgos de selección y composición

En la evaluación de las escuelas y sus efectos y su posterior ordenación se trata de comparar los resultados en términos del rendimiento alcanzado por los estudiantes. En la base de todas las críticas se encuentra la complejidad de lo que encierran las puntuaciones brutas de los tests y de lo que los estudios basados en ellas entienden por efectos de la escuela. Raudenbush y Willms (Raudenbush y Willms, 1995; Willms y Raudenbush, 1989) establecieron una distinción ya clásica entre los diferentes tipos de efectos de las escuelas y lo que representan dentro de las puntuaciones de los tests. Los dos tipos de efectos se basan en el análisis de diferencias entre el rendimiento (o ganancia) de un niño en una escuela particular con el esperado si hubiese estado en otro entorno. La elección del entorno de comparación es crítica para los diferentes usos de la información sobre el efecto de la escuela. Definen dos tipos de efectos de interés, Tipo A y Tipo B. El efecto tipo A es la diferencia entre el rendimiento actual de un alumno y el esperable si hubiese asistido a una escuela «típica», equivalente al que se obtendría si los estudiantes de idénticas características fuesen asignadas al azar a las M escuelas bajo evaluación. Normalmente es el efecto que consideran los padres cuando eligen escuela. Este efecto incluye prácticas de la escuela, composición de los estudiantes y el contexto en el que está localizada. El efecto tipo B se refiere exclusivamente al debido a las prácticas de la escuela, que están bajo la responsabilidad de sus profesionales. Este efecto debe ser separado del contexto de la escuela y otros factores externos (contexto, composición del alumnado, etc.). Este efecto es la diferencia entre el rendimiento de un alumno en una escuela particular y el que sería esperable si asistiese a una escuela con contexto idéntico, con unas prácticas de efectividad promedio. Sería como si M escuelas con contextos idénticos fuesen asignadas a niveles de tratamiento con variación en las prácticas. Este es el efecto *justo* o imparcial que se debería considerar en la evaluación de la eficacia de las escuelas. Una discusión sobre la problemática de establecer inferencias causales en la atribución de dichos efectos a partir de estudios observacionales puede verse en Raudenbush y Willms (1995) y Rubin, Stuart y Zanutto (2004).

Los efectos tipo B se refieren al rendimiento de los estudiantes en una clase y escuela comparado con el de los estudiantes de contextos similares. Estos son los de interés para la rendición de cuentas, ya que otros aspectos no están bajo control de la escuela. La estimación de los efectos tipo A implicaría la sustracción de los efectos de las características de los estudiantes y de las correspondientes interacciones. La

estimación de los efectos tipo B implicaría la sustracción de los efectos de las características de los estudiantes y de los efectos contextuales, junto con las correspondientes interacciones.

Los modelos de VA intentan separar las contribuciones relativas de los distintos efectos, para estimar en la medida de lo posible los efectos de tipo B, atribuibles a la escuela, mediante complejos procedimientos estadísticos que intentan paliar la ausencia de aleatorización en la composición de las escuelas. Se considera que la rendición de cuentas y la evaluación de las escuelas deben limitarse a las partes de la varianza que están bajo su control.

Los modelos antes presentados en la rendición de cuentas y en las tablas de liga originales confunden los efectos de las escuelas con los restantes aspectos que no están bajo su control y que no se deben a las políticas y a las prácticas escolares (Choi, Yamashiro, Seltzer, y Herman, 2004; Rumberger y Palardy, 2004; Linn, 2004, 2005; McCall, Kingsbury y Olson, 2004; Zvoch y Stevens, 2006; Rowe, 2000; Aitkin y Longford, 1986; Hanushek, 1979; Raudenbush, 2004b; Ladd y Walsh, 2002; Novak y Fuller, 2003).

Todos estos efectos derivados de características de los estudiantes, contextos de las escuelas y sus posibles interacciones son los referidos como *sesgos de selección*, ya que la distribución de los estudiantes y de los profesores a las escuelas no es aleatoria. No considerar estos efectos puede llevar a grandes injusticias y a la desmoralización de profesores y directores altamente cualificados y que logran ganancias de aprendizaje considerables con sus estudiantes, pero que atienden a poblaciones en desventaja.

Es necesario realizar los ajustes por las características de los estudiantes y contextuales, lo que lleva a medidas más justas y seguras de los resultados de las escuelas, para que las escuelas puedan ser juzgadas de forma justa (Linn, 2004; 2005). La realización de los ajustes teniendo en cuenta los resultados anteriores de los estudiantes y de otras características, implican necesariamente la utilización de modelos de VA.

## **Funcionamiento y requisitos de los modelos de VA para la evaluación**

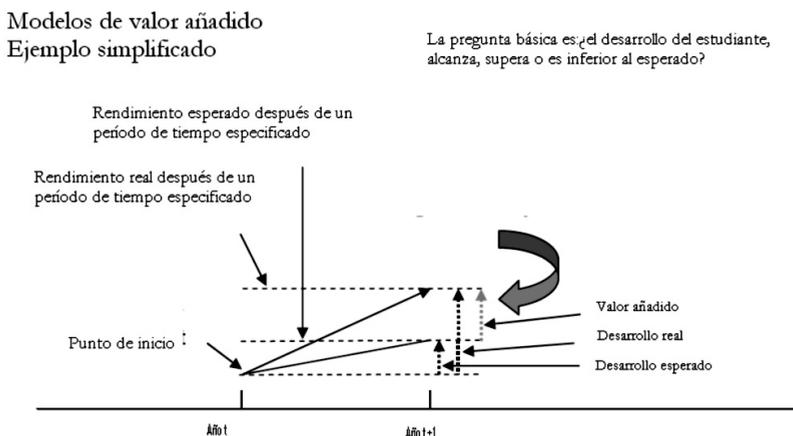
Como se ha visto anteriormente, los modelos de VA son un caso especial de los modelos de desarrollo que intentan aislar los efectos de las escuelas por medio del uso de datos longitudinales (Lissitz, Doran, Schafer y Willhoft, 2006), de modo que las varianzas

en el rendimiento de los estudiantes puedan atribuirse a las escuelas (o profesores en algunos casos). Intentan responder a la cuestión ¿cuánto valor ha añadido la escuela (o el profesor) al aprendizaje del estudiante? Se diferencian de los modelos de desarrollo simple en que intentan atribuir la ganancia o cambio a las escuelas o profesores. En los modelos de VA el término *efecto* representa normalmente la desviación de la escuela de un resultado potencial esperado. Dentro de la terminología de los modelos multinivel referidos más arriba, son los *residuos no explicados* de la escuela después de ajustar o controlar otros factores o fuentes de variabilidad. Los modelos estadísticos utilizados mediante la partición de la varianza intentan aislar los factores diferentes de las prácticas de la escuela (efectos tipo B) para evaluar su influencia en el aprendizaje o desarrollo de los estudiantes (Drury y Doran, 2004; Hershberg, Simon y Lea-Kruger, 2004; McCaffrey et al., 2003).

Los números que estiman el VA son similares a los residuos de la regresión, ya que representan la parte de los resultados (puntuación promedio del estudiante) que no es explicada por las variables explicativas incluidas en el modelo. Al igual que los residuos, estos números tienen de media 0 y el número ligado a cada escuela particular es interpretado provisionalmente como una medida del rendimiento relativo de la escuela. Son estimadores de la diferencia entre la contribución de la escuela al aprendizaje de sus estudiantes y la contribución promedio al aprendizaje de todas las escuelas de las que se obtuvieron datos. Un valor positivo significa que la escuela parece que ha contribuido más al rendimiento que el promedio, es más efectiva, mientras que un valor negativo que es menos efectiva, aunque los estudiantes de la escuela hayan tenido ganancias brutas positivas durante el período bajo estudio. El modelo ajustado y su éxito al explicar los efectos de las escuelas estará determinado por los datos y las variables empleadas en el modelo, así como por el conjunto particular de escuelas de la muestra. Los diferentes estimadores de efectos están sujetos a diferentes tipos de incertidumbre y sesgo, pudiendo establecerse la incertidumbre en sus errores típicos. La ratio del VA estimado respecto de su error típico puede usarse para ver si es estadísticamente diferente de 0 o valor promedio de todas las escuelas.

La lógica implícita o explícita se basa en la comparación de las ganancias de la escuela con otras similares, que comienzan en los mismos niveles de rendimiento o que sirven a poblaciones similares de estudiantes. En la Figura I se presenta una representación simplificada de los modelos de VA.

FIGURA I. Ejemplo simplificado de un modelo de Valor Añadido



## Tipos de modelos estadísticos utilizados en VA

Para la estimación de los efectos de las escuelas pueden utilizarse diferentes modelos estadísticos. La característica fundamental es que incluyen dos o más medidas de rendimiento, pudiendo incluir o no características contextuales de los estudiantes y de las escuelas. Los diferentes modelos parten de distintos supuestos y algoritmos de estimación por lo que los resultados pueden diferir. No obstante, todos tienen en común una serie de características: ser modelos cuantitativos, utilizar las puntuaciones de los estudiantes en tests como medidas de aprendizaje, naturaleza longitudinal con dos o más medidas de los estudiantes e intentar atribuir las ganancias o cambios en desarrollo-aprendizaje a las escuelas o los profesores. Dadas las grandes diferencias entre modelos es difícil unificar la notación. En este apartado seguimos la notación simplificada utilizada en Lissitz et al. (2006). Para un análisis adecuado de la parametrización concreta de los distintos modelos, debe acudir a las fuentes originales.

La primera distinción entre los modelos es la que se establece entre los efectos como *fijos* o *aleatorios*. En la literatura sobre el diseño experimental se consideran efectos fijos aquellos en los que las inferencias se restringen a los términos específicos incluidos en el modelo. Por el contrario, se habla de efectos aleatorios cuando las unidades se consideran una muestra aleatoria de la población e interesa analizar la variabilidad en la población de la que las unidades son muestreadas. En principio, una aproximación no es

necesariamente superior a la otra, sino que está condicionada a las inferencias que se deseen extraer de los datos. La mayor parte de los modelos actuales consideran los efectos de las escuelas como aleatorios, ya que es útil para tratar con estructuras de datos correlacionadas y la heterogeneidad del estatus y cambio de las escuelas.

## Modelos univariantes de efectos fijos

En los casos en los que solamente existen puntuaciones de dos años se puede simplificar la estructura multivariante de los datos y tratar los resultados como univariantes. Hay dos aproximaciones, la denominada de ajuste de covariantes y la de la modelización de las ganancias.

En la aproximación de ajuste de covariantes las puntuaciones del año actual se regresan sobre la puntuación del año anterior y sobre otras covariantes adicionales, si se considera necesario:

$$Y_{it} = \sum_{k=1}^M \beta_k(Y_{t-1,i}) + \sum_{l=1}^J \beta_l(X_{it}) + \varepsilon_{it}; \varepsilon_{it} \sim N(0, \sigma_\varepsilon^2) \quad (8)$$

Donde  $Y_{it}$  representa la puntuación del estudiante  $i$  en el tiempo  $t$ ,  $\beta_k$  es el efecto de la escuela y los  $\beta_l$  representa el coeficiente del covariante  $l$ -ésimo. Este diseño es muy simple, ya que solamente necesita dos puntuaciones de rendimiento que no precisan estar en la misma escala. El valor añadido para el estudiante es la diferencia entre la puntuación actual y la pronosticada y el de la escuela para una materia y curso, el promedio de los residuos de los estudiantes.

Otra aproximación posible es la modelización de la ganancia o de las diferencias entre las puntuaciones de dos años consecutivos de evaluación,  $G_t = Y_t - Y_{t-1}$ , tomándola como la variable dependiente del modelo de regresión:

$$G_{it} = \sum_{k=1}^M \beta_k(x) + \varepsilon_{it}; \varepsilon_{it} \sim N(0, \sigma_\varepsilon^2) \quad (9)$$

Donde  $\beta_k$  es la ganancia promedio de los estudiantes en la escuela  $k$ . Al igual que el modelo anterior es muy fácil de estimar, pero requiere que las dos puntuaciones de los tests estén en una escala comparable. El valor añadido para una escuela y materia es el promedio de los resultados.

Estos modelos de efectos fijos han sido usados preferiblemente en el ámbito de las funciones de producción de la economía de la educación (Lockwood y McCaffrey, 2007).

## Modelos univariantes de efectos aleatorios

Los dos modelos anteriores pueden tratarse asumiendo los efectos de las escuelas como variables aleatorias con modelos de efectos mixtos. Son más frecuentes que los anteriores. Los resultados son similares a los de efectos fijos en cuanto a la variabilidad de las escuelas, pero proporcionan diferentes estimaciones de los efectos que resultan de formas diferentes de tratar con el error muestral. En los modelos de efectos fijos el efecto de la escuela se estima únicamente a partir de los estudiantes de la escuela; en los modelos de efectos aleatorios se utilizan estimadores empíricos de Bayes que contraen el estimador basado en los estudiantes hacia la media global para todos los estudiantes. Aunque sesgados, tienen propiedades estadísticas óptimas excepto para las escuelas cuyos efectos están lejos de la media. Los de efectos fijos están muy afectados por el error muestral cuando las clases o las escuelas son pequeñas.

La extensión del modelo de ajuste de covariante a los efectos aleatorios tendría la forma siguiente, mostrada en la ecuación (6)

$$\begin{aligned} Y_{it} &= \mu + \beta(Y_{t-1,i}) + \theta_{j(i)} + \varepsilon_i \\ \theta_{j(i)} &\sim N(0, \sigma_\theta^2); \varepsilon_i \sim N(0, \sigma_\varepsilon^2) \end{aligned} \quad (10)$$

Donde  $\mu$  es el parámetro media fijado,  $\beta$  es el aumento esperado en  $Y$  asociado con cada cambio unitario en  $Y_{t-1,i}$ ,  $t$  indica el tiempo,  $i$  el estudiante individual y  $\theta$  el efecto de la escuela. La notación  $j(i)$  indica que el estudiante  $i$  está anidado en la escuela  $j$ . Los efectos de la escuela son tratados como variación aleatoria en torno a  $\beta$ .

También se puede reformular el modelo de la ganancia, como en la ecuación 11

$$\begin{aligned} G_i &= \mu + \theta_{j(i)} + \varepsilon_i \\ \theta_{j(i)} &\sim N(0, \sigma_\theta^2); \varepsilon_i \sim N(0, \sigma_\varepsilon^2) \end{aligned} \quad (11)$$

Ahora el efecto de la escuela denota la desviación de la ganancia promedio. En todos los modelos pueden introducirse covariantes adicionales que representen características de los estudiantes.

Puede haber variaciones en los resultados según que se estimen ajustes de covariantes o ganancias (Thum, 2003), ya que puede darse la «paradoja de Lord» (Lord, 1967).

Todos los modelos anteriores comparten algunas limitaciones. Requieren datos completos de los estudiantes, por lo que deben eliminarse sujetos con valores perdidos en alguno de los tests, lo que puede llevar a estimadores sesgados, planteando problemas especiales en escuelas con alta movilidad de los estudiantes y elevadas tasas de repetición. No suelen tener en cuenta de forma explícita los errores de medida, lo que es problemático en los casos de uso de las puntuaciones anteriores como predictores, ya que la regresión asume predictores sin error (McCaffrey et al., 2003). Finalmente, el uso de solamente dos puntuaciones introduce limitaciones en la estimación del desarrollo de los estudiantes (Rogosa, 1995).

Una variante de los modelos anteriores es el desarrollado en el *Dallas Value-Added Assessment System* (Webster y Mendro, 1998; Webster, 2005), que realiza ajustes mediante regresión por mínimos cuadrados ordinarios de las puntuaciones actuales y anteriores, utilizando predictores de características de los estudiantes. Las ganancias o diferencias consideran efectos aleatorios de las escuelas y estos efectos son ajustados en un marco multinivel en el que se consideran de nivel dos y para cuyo ajuste se utilizan variables contextuales de las escuelas.

Otras variaciones incluyen el uso directo de modelos multinivel en los que directamente se modelizan los resultados en un modelo de dos niveles, sin realizar previamente los ajustes de las puntuaciones anteriores y actuales mediante regresión. Los modelos de valor añadido contextualizados de Inglaterra usan este tipo de aproximación.

Los modelos anteriores utilizan como variable dependiente el vector entero de puntuaciones de los sujetos en los tests. Analizan las trayectorias de los estudiantes utilizando para cada sujeto las puntuaciones disponibles de los tests. Suelen incorporar al modelo estructuras complejas de varianza/covarianza de las puntuaciones y permiten incluir diversas trayectorias de desarrollo, tanto lineales como no lineales (Goldstein, 2003; Raudenbush y Bryk, 2002). Pueden darse dos situaciones:

- Modelos completamente anidados, en los que la secuencia de puntuaciones está anidada en los sujetos, y éstos anidados en las clases o en las escuelas, y se denominan modelos anidados de efectos aleatorios, que pueden tratarse con los convencionales modelos multinivel con dos o tres niveles.
- Modelos que permiten analizar diferentes cohortes de estudiantes, la movilidad o cambio de escuelas y el tratamiento de los efectos de escuelas o profesores anteriores.

Existen diferentes procedimientos estadísticos para tratar los modelos anteriores por ejemplo, TVAAS/EVAAS aplica una matriz de covarianza no estructurada a los residuos del estudiante (Ballou, Sanders y Wright, 2004). Otros pueden introducir interceptos y pendientes aleatorias en el nivel del estudiante para explicar esta covarianza (Doran y Lockwood, 2006; Raudenbush y Bryk, 2002).

También existen diseños multivariantes complejos que se suelen diferenciarse en dos bloques, aunque pueden caracterizarse ambos bajo el mismo modelo general (Wright, Sanders y Rivers, 2006).

En primer lugar hay dos modelos cuyo interés radica en la evaluación de los efectos del profesor que son el modelo estratificado (*layered model*) utilizado en los sistemas TVAAS/EVAAS (Ballou et al., 2004; Sanders, 2006; Wright et al., 2006) y el denominado *modelo de la persistencia* (McCaffrey et al., 2003, 2004; Lockwood, 2006). Estos modelos introducen complejidades de cálculo que dificultan su estimación con software convencional.

El segundo bloque, desarrollado en el marco de los modelos multinivel, se conoce como modelo de clasificación cruzada (*cross-classified model*) (Choi, Goldschmidt & Yamashiro, 2006; Choi & Seltzer, 2005; Goldstein, 1987; 2003; Goldstein, Burgess y McConnell, 2007; Ponisciak y Bryk, 2005; Raudenbush y Bryk, 2002).

También se ha propuesto modelos para variables dependientes ordinales (Fielding, Yang y Goldstein, 2003).

Los modelos lineales mixtos utilizados en la mayor parte de las aplicaciones VA son una forma excelente de describir las escuelas efectivas, pero algunos autores encuentran algunos problemas y deficiencias en su utilización, especialmente si se enmarcan dentro de sistemas de rendición de cuentas basadas en estándares como la NCLB. Para responder a la pregunta de si el rendimiento y el desarrollo son adecuados, se han propuesto algunos modelos que están recibiendo cierta consideración en la literatura especializada: sistema REACH (Doran e Izumi, 2004), el modelo híbrido (Kingsbury y McCall, 2006; McCall, Kingsbury y Olson, 2004), las tablas de resultados (Hill et al, 2006) y las matrices de transición (Betebenner, 2005).

## Conclusiones

Los modelos de VA son un conjunto de procedimientos estadísticos que han despertado un gran interés entre los investigadores en educación, administradores escolares

y políticos, ya que proporcionan medios para separar los efectos derivados de las prácticas de las escuelas de variables de los estudiantes y contextuales. Se consideran un desarrollo muy prometedor para la identificación de las buenas escuelas y de las que necesitan ayuda. Su aplicación mide los efectos de las escuelas (y de los profesores) con mayor precisión y justicia que otros procedimientos utilizados con frecuencia en la evaluación, ya que en el aprendizaje de los estudiantes intervienen numerosos factores, tanto escolares como no escolares.

Durante los últimos diez años se han hecho importantes avances en el desarrollo de estos modelos. El uso de métodos estadísticos complejos ha permitido seguir las trayectorias de rendimiento de los estudiantes con mediciones longitudinales múltiples y la incorporación de medidas de ajuste para separar los efectos verdaderos de la escuela en las ganancias de aprendizaje de los estudiantes.

Los resultados se han mostrado útiles en varias actividades: rendición de cuentas, con o sin consecuencias para las escuelas, elección de escuelas basada en resultados más objetivos que las tablas de liga, y uso de la información en los procesos de mejora y desarrollo de las escuelas, con importantes aplicaciones para los procesos de autoevaluación. En general, sus datos pueden considerarse una buena fuente para el diálogo con las escuelas y la reflexión para emprender acciones de mejora, que mejorarán el rendimiento de los estudiantes.

No obstante, aunque parece haber ganado una extraordinaria popularidad, algunas características metodológicas y de su aplicación en la práctica están todavía bajo escrutinio. En Martínez Arias (2008, en este volumen) se presenta una revisión del estado del arte y de la problemática planteada por los modelos. Cuestiones sobre los supuestos, el tratamiento de los valores perdidos, las diferencias derivadas del uso de uno u otro modelo, la incertidumbre de los estimadores, la dificultad de comunicar los resultados a las partes interesadas de forma comprensible, la necesidad de disponer de tests que permitan la presentación de los resultados en una escala única y la atribución de los efectos a las escuelas a partir de datos observacionales y no experimentales, así como su carácter normativo, que dificulta establecer valoraciones de si el aprendizaje es adecuado, junto con la posibilidad de que rebajen las expectativas para ciertos grupos de sujetos, están bajo discusión en la actualidad.

A pesar de todo, las investigaciones recientes están ayudando a la superación de algunos de los anteriores problemas y los resultados combinados con procesos sistemáticos para la interpretación y aplicación de los datos, pueden ayudar a políticos, administradores, inspectores, directores de centros y profesores a proporcionar una educación de calidad para todos los estudiantes.

## Presentación del monográfico

Como hemos presentado a lo largo de estas páginas, los modelos de valor añadido, más que una metodología estadística para tratar los datos procedentes de la evaluación educativa, constituyen una nueva forma de entender el papel que la evaluación desempeña o puede desempeñar en el seno de los sistemas educativos.

Es muy difícil exagerar la importancia que esta estrategia evaluativa va a adquirir en los próximos años en el funcionamiento ordinario de la educación, aunque hoy todavía se trate fundamentalmente de un tema de investigación, al que se incorporan destacadas y bien conocidas experiencias aplicadas.

Todos estos temas y los problemas prácticos y metodológicos asociados son tratados en este número monográfico. En él encontramos diez contribuciones en las que se dibuja el amplio y complejo panorama del valor añadido en educación. Se pueden agrupar en tres grandes bloques temáticos, que muestran la lógica interna de los modelos de valor añadido y la relevancia y vigencia de los mismos.

El primer bloque temático incluye dos artículos. El primero de éstos de los profesores Rosario Martínez Arias, José Luis Gaviria y María Castro (Universidad Complutense de Madrid), se centra en la conceptualización y evolución de los distintos modelos de valor añadido. No se podía comenzar este monográfico sin una adecuada definición de esta medida, mostrando que su dimensión teórica no es ajena a las distintas aproximaciones a la evaluación de la contribución específica de las escuelas a los aprendizajes escolares. De ahí que en esta aportación, además de mostrar los principales modelos de valor añadido en uso, también se repasan los principales retos y dificultades conceptuales y metodológicas de su medición.

Muy vinculado a este artículo, la contribución de la profesora Martínez Arias en el artículo *Usos, aplicaciones y problemas de los modelos de valor añadido en educación* tiene un carácter más metodológico, pues se revisan los modelos estadísticos en uso y se presenta una revisión de los problemas estadísticos, psicométricos y prácticos relacionados con el uso de los modelos de valor añadido para favorecer una interpretación prudente y transparente de los mismos.

El segundo bloque temático se dedica a la descripción y análisis de algunas de las experiencias prácticas vigentes de aplicación de modelos de evaluación de sistemas basados en el valor añadido. Los tres siguientes artículos muestran los sistemas de valor añadido inglés y norteamericano que representan los ejemplos reales de aplicación de estos modelos.

El artículo de Ray, Evans y McCormack muestra la experiencia inglesa desarrollada desde el año 2002 hasta la actualidad con todas las escuelas primarias y secundarias.

El punto clave de su medición del valor añadido es el desarrollo de una medida sencilla a la vez que técnicamente precisa que es utilizada tanto por las escuelas para el desarrollo de sus propios planes de mejora, como por la Administración educativa inglesa para la introducción y seguimiento de cambios en el sistema educativo que ayuden a elevar los estándares de rendimiento.

El artículo del profesor Thum de la Universidad de Michigan State (USA) recoge una reflexión crítica y profunda, así como un conjunto de pautas de carácter práctico para y sobre el desarrollo de los modelos de valor añadido que pueden acogerse dentro de la legislación norteamericana *No Child Left Behind*, que sin duda ha supuesto una revolución y un estímulo para estos modelos desde 2002.

El último bloque conceptual está relacionado con las *cuestiones metodológicas*, importantes si se quiere desarrollar un sistema de rendición de cuentas basado en modelos de valor añadido, e incluye seis artículos. Si bien es cierto que existe consenso en la literatura de investigación sobre la definición y relevancia práctica de los sistemas de rendición de cuentas de los sistemas educativos, tampoco hay duda de que la discusión científica de los modelos de valor añadido se centra en cómo traducir operativamente las ideas sobre lo que debería ser el valor añadido en la práctica, de forma que se somete a constante escrutinio y mejora el mecanismo técnico subyacente a estos modelos. Sin una adecuada especificación metodológica, no se puede establecer la verdadera utilidad de estos modelos.

El artículo de Eva L. Baker, se centra en la validez de los sistemas de evaluación usados en las escuelas, bien sean éstos formativos o de rendición de cuentas. El inicio de los sistemas de evaluación está en la medida de los logros académicos; de ahí que la autora demande que los tests tengan la denominada *sensibilidad instructiva*.

Vinculados también a cuestiones de medida del rendimiento académico, se incluyen otros dos artículos, relacionados con la necesaria unidimensionalidad de las pruebas para la medida a través del tiempo del valor añadido.

El artículo de los profesores Lizasoain y Joaristi realiza un análisis de la dimensionalidad de un conjunto de pruebas de Matemáticas utilizadas en una evaluación longitudinal del valor añadido realizada con alumnos desde 5º de Educación Primaria hasta 4º de Educación Secundaria Obligatoria de la Comunidad de Madrid, de 2005 a 2007. La clave de este trabajo está en la triangulación de los resultados obtenidos por técnicas factoriales clásicas, por los procedimientos de análisis factorial de información total y por métodos no paramétricos basados en la Teoría de Respuesta al Ítem.

Muy unido a esta temática está el artículo de González, Blanco y Ordóñez. En él, se parte de los patrones de correlaciones que presentan las sucesivas medidas de

rendimiento necesarias para obtener una medida de valor añadido. Cuando se comparan estos patrones con los que aparecen cuando se correlacionan las variables estructurales, se comprueba cómo la fiabilidad y la complejidad estructural de las medidas de rendimiento afectan a esos patrones observables. Una conclusión importante de este artículo es que la creciente complejidad que presentan las medidas de rendimiento académico de cursos sucesivos es más importante, a la hora de explicar los patrones de correlación, que la propia fiabilidad de los instrumentos utilizados.

El concepto de valor añadido y el concepto de crecimiento y cambio en los aprendizajes de los alumnos están íntimamente relacionados. El artículo de Castro, Ruiz de Miguel y López, estudia los efectos del principal predictor de la tasa de crecimiento individual: el nivel inicial de conocimientos de los alumnos. La inclusión de este predictor permite el control de artefactos estadísticos no deseados como la regresión estadística. También se estudian distintas formas de crecimiento distintas a la más utilizada clásicamente que es la lineal. De esta manera, se puede establecer cuál es la forma básica de la función de crecimiento en los modelos de valor añadido.

El artículo de la profesora Ferraõ, de la Universidad da Beira Interior (Portugal), describe el grado en que dos medidas diferentes del estatus socioeconómico, una variable clásica en los modelos de valor añadido, o incluso la exclusión de una variable, producen cambios importantes en las estimaciones del valor añadido de cada escuela.

Por último, el artículo escrito por los profesores Gaviria, Biencinto y Navarro estudia la posibilidad de mantener la misma estructura longitudinal de base con los alumnos de Primaria y Secundaria. La relevancia de este artículo estriba en que es en la transición entre estas dos etapas donde se identifican las mayores variaciones en los indicadores del logro académico, y una estructura común permitiría un análisis de los factores diferenciales que afectan al rendimiento desde la misma estructura de varianza-covarianza.

El conjunto de estos diez artículos representa desde nuestro punto de vista un repaso en profundidad a las cuestiones más candentes relacionadas con el concepto, la medida, los modelos, la metodología y las experiencias aplicadas en el ámbito del valor añadido. Nuestra intención como coordinadores de este volumen ha sido ofrecer al lector un balcón privilegiado y actualizado desde el cual asomarse al complejo y apasionante mundo de la evaluación de los sistemas educativos mediante el uso de modelos de valor añadido. Esperamos que este acercamiento pueda satisfacer y estimular la curiosidad y las diversas necesidades de los lectores de la Revista de Educación.

## Referencias bibliográficas

- AITKIN, M. & LONGFORD, N. (1986). Statistical modeling issues in school effectiveness studies. *Journal of the Royal Statistical Society, Series A*, 149, 1-43.
- BALLOU, D., SANDERS, W. & WRIGHT, P. (2004). Controlling for student background in value-added assessment of teachers. *Journal of Educational and Behavioral Statistics*, 29, 37-65.
- BETEBENNER, D.W. (2005). *Performance standards in measures of educational effectiveness*. Boston College. Department of Educational Research, Measurement and Evaluation.
- BRAUN, H. & KANJEE, A. (2006). *Using assessment to improve education in developing nations*. En H. BRAUN, A. KANJEE, BETTINGER, R. & KREMER, M. (eds.), *Improving education through assessment, innovation, and evaluation* (1-46). Cambridge, MA: American Academy of Arts and Sciences.
- BURNSTEIN, L. (1980). The analysis of multilevel data in educational research and evaluation. *Review of Research in Education*, 8, 158-233.
- CAREY, K. (2004). The real value of teachers: Using new information about teacher effectiveness to close achievement gap. *Thinking K-16*, 8, 1-42.
- CARLSON, K., MARTÍNEZ, F., O'DAY, J., STECHER, B., TAYLOR, J. & COOK, A. (2007). *State and local implementation of the No Child Left Behind Act. Vol. III. Accountability under NCLB Interim report*. Santa Mónica, CA: The RAND Corporation.
- CHOI, K., GOLDSCHMIDT, P. & YAMASHIRO, K. (2006). *Exploring models of school performance: From theory to practice* (CSE Rep: No. 673). Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing.
- CHOI, K. & SELTZER, M. (2005). *Modeling Heterogeneity in Relationships between Initial Status and Rates of Change: Latent Variable Regression in a Three-Level Hierarchical Model*. (CSE Rep. No 647). Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing.
- CHOI, K., SELTZER, M., HERMAN, J. & YAMASHIRO, K. (2004). *Children left behind in AYP and non-AYP schools: Using student progress and the distribution of student gains to validate AYP* (CSE Rep. No 637). Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing.
- CIZEK, G. J. & BUNCH, M. B. (2007). *Standard setting: a guide to establishing and evaluating performance standards on tests*. Thousand Oaks, CA: Sage.
- COLEMAN, J. S., CAMPBELL, E. Q., HOBSON, C. J., MCPARTLAND, J., MOOD, A. M., WEINFELD, F. D. & YORK, R. L. (1966). *Equality of educational opportunity*. Washington, DC: U.S. Government Printing Office.

- CRONIN, J., KINGSBURY, G. G., MCCALL, M. S. & BOWE, B. (2005). *The impact of No Child Left Behind act on student achievement and growth*. Portland, OR: Northwest Evaluation Association.
- DE LEEUW, J. & MEIJER, E. (2008a). Introduction to multilevel analysis. En J. DE LEEUW & E. MEIJER (eds.), *Handbook of multilevel analysis* (pp. 1-75). New York: Springer.
- (2008b). *Handbook of multilevel analysis*. New York: Springer.
- DORAN, H. & LOCKWOOD, J. R. (2006). Fitting value-added models in R. *Journal of Educational and Behavioral Statistics*, 31, 205-230.
- DORAN, H. C. & IZUMI, L. T. (2004). *Putting Education to the Test: A Value-Added Model for California*. San Francisco, CA: Pacific Research Institute.
- DRURY, D. & DORAN, H. (2003). The Value of Value-Added Analysis. *NSBA Policy Research Brief*, 3, 1-4.
- EDMONDS, R. R. (1979). Effective schools for the urban poor. *Educational Leadership*, 37, 15-27.
- FIELDING, A., YANG, M. & GOLDSTEIN, H. (2003). Multilevel ordinal models for examination grades. *Statistical Modelling*, 3, 127-153.
- FITZ-GIBBON, C. T. (1997). *The value-added national project: Final report: feasibility studies for a national system of value added indicators*. London: School Curriculum and Assessment Authority.
- GAVIRIA, J. L. Y CASTRO, M. (2005). *Los modelos jerárquicos lineales*. Madrid: La Muralla.
- GELMAN, A. & HILL, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. New York: Cambridge University Press.
- GOLDSCHMIDT, P., CHOI, K. & MARTINEZ, F. (2004). *Using Hierarchical Growth Models To Monitor School Performance Over Time: Comparing NCE to Scale Score Results (CSE Rep. No 618)*. Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing.
- GOLDSTEIN, H. & SPIEGELHALTER, D. (1996). League Tables and Their Limitations: Statistical Issues in Comparisons of Institutional Performance. *Journal of the Royal Statistical Society, A*, 159, 385-443.
- GOLDSTEIN, H. & THOMAS, S. (1996). Using Examination Results as Indicators of School and College Performance. *Journal of Royal Statistical Society, A*, 159, 149-163.
- GOLDSTEIN, H., RASBASH, J., YANG, M., WOODHOUSE, G., PAN, H., NUTTALL, D. & THOMAS, S. (1993). A multilevel analysis of school examination results. *Oxford Review of Education*, 19, 425-433.
- GOLDSTEIN, H. (1987a). *Multilevel models in educational and social research*. New York: Oxford University Press.

- (2003b). *Multilevel models*. London: Arnold.
- GOLDSTEIN, H., BURGESS, S. & MCCONNELL, B. (2007). Modelling the effect of pupil mobility on school differences in educational achievement. *Journal of The Royal Statistical Society, A*, 170, 941-954.
- GRAY, J., JESSON, D., GOLDSTEIN, H., HEDGER, K. & RASBASH, J. (1995). A multilevel analysis of school improvement: changes in schools' performance over time. *School Effectiveness and School Improvement*, 10, 97-114.
- HAMILTON, L. (2003). Assessment as a policy tool. *Review of Research in Education*, 27, 25-68.
- HANUSHEK, E. A. (1979a). Conceptual and empirical issues in the estimation of educational production functions. *Journal of Human Resources*, 14, 351-388.
- (2003). The failure of input-based schooling policies. *Economic Journal*, 113, 64-98.
- (2005). The economics of school quality. *German Economic Review*, 6, 269-286.
- HANUSHEK, E. A. & RAYMOND, M. E. (2004). The Effect of School Accountability Systems on the Level and Distribution of Student Achievement. *Journal of the European Economic Association*, 2, 406-415.
- HANUSHEK, E. A. & WOSSMANN, L. (2006). Does Early Tracking Affect Educational inequality and performance? Differences-in-differences evidence across countries. *Economic Journal*, 116, 63-76.
- HILL, R. & DEPASCALE, C. (2003). Reliability of No Child Left Behind accountability designs. *Educational Measurement: Issues and Practices*, 22(3), 12-20.
- HILL, R., SCOTT, M., DE PASCALE, C., DUNN, J. & SIMPSON, M. A. (2006). Using value tables to explicitly value student growth. En R. LISSITZ (ed.), *Longitudinal and value added models of student performance* (255-290). Maple Grove, MN: JAM Press.
- HOX, J. (2002). *Multilevel analysis: Techniques and applications*. Mahwah, NJ: Erlbaum.
- JENCKS, C., SMITH, M. S., ACKLAND, H., BANE, M. J., COHEN, D., GRINTLIS, H., HEYNES, B. & MICHELSON, S. (1972). *Inequality: A reassessment of the effect of family and schooling in America*. New York: Basic Books.
- KANE, T. J. & STAIGER, D. O. (2002). The promise and pitfalls of using imprecise school accountability measures. *Journal of Economic Perspectives*, 16, 91-114.
- KINGSBURY, G. G. & MCCALL, M. S. (2006). *The hybrid success model: Theory and practice*. En R. LISSITZ (ed.), *Longitudinal and value added models of student performance* (346-379). Maple Grove, MN: JAM Press.
- LADD, H. F. & WALSH, R. P. (2002). Implementing value-added measures of school effectiveness: Getting the incentives right. *Economics of Education Review*, 21, 1-17.

- LINN, R. L., BAKER, E. V. & BETEBENNER, D. W. (s.f.). Accountability systems: Implications of requirements of the no child left behind act of 2001. *Educational Researcher*, 31, 3-16.
- LINN, R. L. & HAUG, C. (2002). Stability of school building accountability scores and gains. *Educational Evaluation and Policy Analysis*, 24, 29-36.
- LINN, R. L. (2004). Assessment and accountability. *Educational Researcher*, 29, 4-14.
- LISSITZ, R., DORAN, H., SCHAFER, W. & WILLHOFT, J. (2006). Growth modeling, value-added modeling and linking: An introduction. En R. LISSITZ (ed.), *Longitudinal and Value-Added Models of Student Performance* (pp. 1-46). Maple Grove, Minnesota: JAM Press.
- LOCKWOOD, J. R. (2006). A case study of some practical challenges of longitudinal student achievement modeling: The RAND Mosaic II Study. En R. LISSITZ (ed.), *Longitudinal and Value-Added Models of Student Performance* (230-254). Maple Grove, Minnesota: JAM Press.
- LOCKWOOD, J. R. & MCCAFFREY, D. F. (2007). Controlling for individual heterogeneity in longitudinal models with applications to student achievement. *Electronic Journal of Statistics*, 1, 223-252.
- LORD, F. M. (1967). A paradox in the interpretation of group comparisons. *Psychological Bulletin*, 68, 304-305.
- MACBEATH, J. & MORTIMORE, P. (2001). *Improving school effectiveness*. Buckingham: Open University Press.
- MARCHESI, A. Y MARTÍNEZ ARIAS, R. (2006). *Escuelas de éxito en España. Sugerencias e interrogantes a partir del informe PISA 2003*. Madrid: Fundación Santillana.
- MARTÍNEZ ARIAS, R., HERNÁNDEZ LLOREDA, V. Y HERNÁNDEZ LLOREDA, M. J. (2006). *Psicometría*. Madrid: Alianza.
- MCCAFFREY, D. F., LOCKWOOD, J. R., KORETZ, D. M. & HAMILTON, L. S. (2003). *Evaluating Value-Added Models for Teacher Accountability*. Santa Mónica, CA: The RAND Corporation.
- MCCAFFREY, D. M., LOCKWOOD, J. R., KORETZ, D., LOUIS, T. A. & HAMILTON, L. (2004). Models for value-added modeling of teacher effects. *Journal of Educational and Behavioral Statistics*, 29, 67-101.
- MCCALL, M. S., KINGSBURY, G. G. & OLSON, A. (2004). *Individual growth and school success*. Portland, OR: Northwest Evaluation Association.
- MORTIMORE, P., SAMMONS, P. & THOMAS, S. (1994) School Effectiveness and Value Added Measures, *Assessment in Education: Principles, Policy y Practice*, 1, 315-332.
- MORTIMORE, P., SAMMONS, P., STOLL, L., LEWIS, D. & ECOB, R. (1988). The effects of school membership on pupils' educational outcomes. *Research Papers in Education*, 3(1), 3-26.

- MUIJS, R. D. & REYNOLDS, D. (2000). School effectiveness and teacher effectiveness: some preliminary findings from the evaluation of the Mathematics Enhancement Programme. *School Effectiveness and School Improvement*, 11, 273-303.
- NO CHILD LEFT BEHIND ACT OF 2001 (2002). Public Law No. 107-110, 115 Stat. 1425.
- NUTTALL, D. L., GOLDSTEIN, H., PROSSER, R. & RASBASH, J. (1989). Differential School Effectiveness. *International Journal of Educational Research*, 13, 769-776
- O'DAY, J. A. & SMITH, M. S. (1993). *Systemic reform and educational opportunity*. En S. H. Fuhrman (ed.), *Designing coherent education policy: Improving the system* (250-312). San Francisco: Jossey Bass.
- PONISCIAK, S. M. & BRYK, A. S. (2005). Value-added analysis of the Chicago public schools: An application of hierarchical models. En R. Lissitz (ed.), *Value-Added models in education: Theory and applications* (pp. 40-79). Mapple Grove, MN: JAM Press.
- RAUDENBUSH, S. (2004). What are value-added models estimating and what does this imply for statistical practice? *Journal of Educational and Behavioral Statistics*, 29, 121-129.
- RAUDENBUSH, S. W. & BRYK, A. S. (1986). A hierarchical model for studying school effects. *Sociology of Education*, 59, 1-17.
- (2002). *Hierarchical Linear Models* (2ª ed.). Thousand Oaks, CA: Sage.
- RAUDENBUSH, S. W. & WILLMS, J. D. (1995). The estimation of school effects. *Journal of Educational and Behavioral Statistics*, 20, 307-335.
- RAY, A., EVANS, H. & McCORMACK, T. (2008). The use of national value added models for school improvement in English schools. *Revista de Educación*, 348.
- REYNOLDS, D. & CREEMERS, B. (1990). School Effectiveness and School Improvement: a Mission Statement. *School Effectiveness and School Improvement*, 1, 1-3
- ROBINSON, W. (1950). Ecological correlations and the behavior of individuals. *American Sociological Review*, 15, 351-357.
- ROGOSA, D. (1995). *Myths about longitudinal research*. En J. M. Gottman (ed.), *The analysis of change* (3-66). Mahwa, NJ: Erlbaum.
- ROWE, K. J. (2000). Assessment, league tables and school effectiveness: Consider the issues and «Let's get real». *Journal of Educational Enquiry*, 1, 73-98.
- RUBIN, D. B., STUART, E. A. & ZANUTTO, E. E. (2004). A potential outcomes view of value-added assessment in education. *Journal of Educational and behavioural Statistics*, 29, 103-116.
- RUMBERGER, R. W. & PALARDY, G. J. (2003). *Multilevel models for school effectiveness research*. En D. KAPLAN (ed.), *Handbook of Quantitative Methodology for the Social Sciences* (235-258). Thousand Oaks, CA: Sage.

- SAMMONS, P. (1996). Complexities in the judgement of school effectiveness. *Educational Research and Evaluation*, 2, 113-49.
- SAMMONS, P., HILLMAN, J. & MORTIMORE, P. (1995). *Key Characteristics of Effective Schools: A review of school effectiveness research*. London: Office for Standards in Education.
- SAMMONS, P. & REYNOLDS, D. (1997). A partisan Evaluation John Elliot of school effectiveness. *Cambridge Journal of Education*, 27, 123-36.
- SANDERS, W. & HORN, S. (1994). The Tennessee value-added assessment system (TVAAS): Mixed model methodology in educational assessment. *Journal of Personnel Evaluation*, 9, 299-311.
- (1998). Research findings from the Tennessee value-added assessment system (TVAAS) database: Implications for educational evaluation and research. *Journal of Personnel Evaluation in Education*, 12, 247-256.
- SANDERS, W. L., SAXTON, A. M. & HORN, S. P. (1997). The Tennessee value-added assessment system: a quantitative, outcomes-based approach to educational assessment. En J. MILLMAN (ed.), *Grading teachers, grading schools: Is student achievement a valid evaluation measure?* (pp. 137-162). Thousand Oaks, CA: Corwin Press.
- SAUNDERS, L. (1999). A brief history of educational «Value Added»: How did we get to where we are? *School Effectiveness and School Improvement*, 10, 233-256.
- SCHAGEN, I. (2006). The use of standardized residuals to derive value-added measures of school performance. *Educational Studies*, 32, 119-32.
- SCHEERENS, J. (2005). The quality imperative. Paper commissioned for the EFA Global Monitoring report 2005. *Review of school and instructional effectiveness research*.
- SCHEERENS, J. & BOSKER, R. J. (1997). *The Foundations of Educational Effectiveness*. Oxford: Elsevier.
- STEVENS, J. (2005). *The study of school effectiveness as a problem in research design*. En R. LISSITZ (ed.), *Value-Added models in education: Theory and applications* (pp. 166-208). Maple Grove, MN: JAM Press.
- TAYLOR, J. & NGUYEN, A. N. (2006). An Analysis of the Value Added by Secondary Schools in England: Is the Value Added Indicator of Any Value? *Oxford Bulletin of Economics y Statistics*, 68, 203-224.
- TEDDLIE, C. & REYNOLDS, D. (2000). *The International Handbook of School Effectiveness Research*. New York: Falmer Press.
- THUM, Y. M. (2002). *Measuring student and school progress with the California API*. Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing.

- (2003). Measuring progress toward a goal: estimating teacher productivity using a multivariate multilevel model for value-added analysis. *Sociological Methods and Research*, 32, 153-207.
- WANG, L., BECKETT, G. H. & BROWN, L. (2006). Controversies of standardized assessment in school accountability reform: A critical synthesis of multidisciplinary research evidence. *Applied Measurement in Education*, 19, 305-328.
- WEBSTER, W. J. & MENDRO, R. L. (1997). The Dallas value-added accountability system. En J. MILLMAN (ed.), *Grading teachers, grading schools: Is student achievement a valid evaluation measure?* (pp. 81-99). Thousand Oaks, CA: Corwin Press.
- WEBSTER, W. J. (2005). The Dallas school level accountability model: The marriage of status and value-added approaches. En R. LISSITZ (ed.), *Value added models in education: Theory and applications* (pp. 233-271). Maple Grove, MN: JAM Press.
- WILLMS, J.D. & RAUDENBUSH, S.W. (1989). A longitudinal hierarchical linear model for estimating school effects and their stability. *Journal of Educational Measurement*, 2, 209-232.
- WRIGHT, S. P., SANDERS, W. L. & RIVERS, J. C. (2006). *Measurement of academic growth of individual students toward variable and meaningful academic standards*. En R. LISSITZ (ed.), *Longitudinal and value added models of student performance* (385-406). Maple Grove, MN: JAM Press.
- YANG, M., GOLDSTEIN, H., RATH, T. & HILL, N. (1999). The use of assessment data for school improvement purposes. *Oxford Review of Education*, 25, 469-483.

## Fuentes electrónicas

- GOLDSCHMIDT, P., ROSCHEWSKI, P. CHOI, K., AUTY, W., HEBBLER, S., BLANK, R. & WILLIAMS, A. (2005). *Policymakers' guide to growth models for school accountability: How do accountability models differ?* Washington, DC: Councils of Chief State School Officers, de <http://www.ccsso.org/publications/>
- HERSHBERG, T., SIMON, V. A. & LEA-KRUGER, B. (February 2004). Measuring What Matters. *American School Board Journal*. National School Boards Association. Consultado el 25 de octubre de 2007, de <http://www.asbj.com/2004/02/0204asbjhershberg.pdf>
- LINN, R. L. (2005). Conflicting demands of No Child Left Behind and state systems: Mixed messages about school performance. *Education Policy Analysis Archives*, 13(33). Consultado el 30 de octubre de 2006, de <http://epaa.asu.edu/epaa/v13n33/>

- NATIONAL COMMISSION OF EXCELLENCE IN EDUCATION (1983). *A Nation at risk*. Washington, DC: Author. Consultado de <http://www.ed.gov/pubs/NatAtRisk/index.html>
- RAUDENBUSH, S. (2004a). *Schooling, statistics, and poverty: Can we measure school improvement?* Princeton, NJ: Educational Testing service. Consultado de <http://www.ets.org/research/pic/angoff9.pdf>
- SANDERS, W.L. (2006). *Comparisons Among Various Educational Assessment Value-Added Models*. The Power of Two-National Value-Added Conference, Columbus, OH. Consultado el 14 de junio de 2007, de <http://www.sas.com/govedu/edu/services/vaconferencepaper.pdf>
- ZVOCH, K. & STEVENS, J. J. (2003). A multilevel, longitudinal analysis of middle school math and language achievement. *Education Policy Analysis Archives*, 11(20). Consultado el 20 de octubre de 2007, de <http://epaa.asu.edu/epaa/v11n20>
- (2006). Successive student cohorts and longitudinal growth models: An investigation of elementary school mathematics performance. *Education Policy Analysis Archives*, 14(2). Consultado el 20 de octubre de 2007, de <http://epaa.asu.edu/epaa/v14n2/>

**Dirección de contacto:** Rosario Martínez Arias. Universidad Complutense de Madrid. Departamento de Metodología de las Ciencias del Comportamiento. Facultad de Psicología. Campus de Somosaguas, 28223 Pozuelo de Alarcón, Madrid. E-mail: [rmnez.arias@psi.ucm.es](mailto:rmnez.arias@psi.ucm.es)



# El uso de los modelos nacionales de valor añadido para la mejora de las escuelas británicas

## The use of national value-added models for school improvement in English schools

Andrew Ray, Helen Evans, Tanya McCormack

*Department for Children, Schools and Families, London, United Kingdom*

### Resumen

La mejora de la escuela es central en la política educativa en Inglaterra. Se anima a las escuelas a desarrollarse y mejorar dentro de un marco de autoevaluación, inspección, consejo y asesoramiento en el nivel local y nacional, y a tomar parte en distintas iniciativas y políticas. La recogida y uso de datos es crucial en este proceso. El análisis de datos tiene lugar en todos los niveles del sistema, desde la evaluación de políticas nacionales, hasta el nivel de alumnos individuales, donde las escuelas usan los datos para seguir el progreso académico y la evaluación de los aprendizajes. En las escuelas británicas se han desarrollado un abanico de iniciativas para tratar de elevar los estándares de rendimiento. Una pieza fundamental ha sido el desarrollo de pruebas que pueden vincularse a los logros previos de los estudiantes y usarse para generar estimaciones de la contribución de la escuela al progreso de sus alumnos: indicadores de valor añadido. En Inglaterra, la disponibilidad de datos nacionales de las características de los estudiantes individuales desde 2002 ha permitido el desarrollo de puntuaciones «contextuales de valor añadido» para escuelas y grupos de alumnos dentro de cada escuela. Estas características pueden analizarse con un paquete estadístico llamado RAISEonline, desarrollado conjuntamente por el gobierno central y la agencia independiente de inspección de las escuelas. Este artículo describe estos desarrollos y aporta ejemplos del tipo de análisis que RAISEonline puede realizar u ofrecer.

*Palabras clave:* modelos de valor añadido, sistema educativo inglés, sistemas de evaluación, sistemas de rendición de cuentas, modelos de crecimiento.

## Abstract

School improvement is central to education policy in England. Schools are encouraged to develop and improve within a framework of self-evaluation, inspection, advice and guidance at local and national level, and by taking part in various initiatives and policies. The collection and use of data is crucial in this process. Analysis of data takes place at all levels of the system, ranging from the evaluation of national policies, down to the level of individual pupils, where schools use data in tracking progress and assessment for learning. A range of school improvement initiatives have been developed to try to raise standards in English schools. A crucial building block has been the development of test data which can be matched to pupils' prior attainment and used to generate estimates of schools' contribution to the progress of their pupils: value-added indicators. In England, the availability of national data on the characteristics of individual pupils since 2002 has allowed the development of «contextual value-added» scores for schools and groups of pupils within schools. These statistics can be analysed in a software package called RAISEonline, developed jointly by central government and the independent school inspection agency. The paper describes these developments and provides examples of the kind of analysis RAISEonline can now provide.

*Key Words:* Value-added models, English educational system, assessment system, accountability system, growth models.

## Introducción

La mejora de la escuela es central en la política educativa en Inglaterra. Se anima a las escuelas a desarrollarse y mejorar dentro de un marco de autoevaluación, inspección, consejo y asesoramiento en el nivel local y nacional, y a tomar parte en distintas iniciativas y políticas. La recogida y uso de datos es crucial en este proceso. El análisis de datos tiene lugar en todos los niveles del sistema, desde la evaluación de políticas nacionales, hasta el nivel de alumnos individuales, donde las escuelas usan los datos para seguir el progreso académico y la evaluación de los aprendizajes.

En Inglaterra hay aproximadamente 8,1 millones de estudiantes en 25.000 escuelas de financiación pública e independientes (DfES 2007) -el 7% de los alumnos están en la enseñanza independiente-. Algunos estudiantes con necesidades educativas especiales están escolarizados en escuelas regulares, otros están escolarizados en escuelas específicas de forma separada. Hay aproximadamente 17.400 escuelas primarias, que generalmente acogen alumnos desde los 4 a los 11 años, y alrededor de 3.300

escuelas secundarias, que normalmente cubren las edades de 11 a 16 años (1.750 tienen «seis formas» que cubren también más allá de los 16 años). El tamaño medio de una escuela secundaria es de 980 estudiantes (aproximadamente 140 estudiantes por cohorte anual en términos promedios); las escuelas primarias tienen en promedio 240 alumnos, 40 por año.

El mantenimiento de las escuelas se realiza a través del gobierno local: hay 150 autoridades locales en toda Inglaterra. Las autoridades locales varían considerablemente en tamaño y características. La menor es la de Scilly Isles, con únicamente una escuela, y la mayor es Kent con 103 escuelas secundarias y 466 escuelas primarias. Los datos educativos son usados extensivamente en el nivel de la autoridad local y en el más amplio nivel regional (por ejemplo, para observar el progreso de los alumnos de Londres). Los análisis son realizados por redes o grupos geográficos de escuelas, por las propias escuelas y dentro de las escuelas por áreas temáticas, profesores y por grupos de estudiantes con características específicas. Este artículo estudia algunos de los análisis desarrollados centralmente con datos recogidos de los alumnos individuales y de las escuelas (no hay una recogida de datos nacional que vincule a los profesores con logros de los alumnos). Se centra en los recientes desarrollos en el uso de modelos de valor añadido y en indicadores para la mejora escolar.

El punto de partida para cualquier análisis de mejora de las escuelas es el desarrollo de pruebas fiables y consistentes o datos de evaluación. El actual sistema en Inglaterra para alumnos de hasta 16 años implica evaluaciones al final de las cuatro etapas clave. Estas etapas están definidas como parte del Currículo Nacional que, introducido en 1988, establece las materias y los programas de estudio que las escuelas públicas están obligadas a cubrir. Las etapas claves 1, 2, 3 y 4 cubren los rangos de edad de 5 a 7 años, de 7 a 11 años, de 11 a 14 años y de 14 a 16 años respectivamente. Había pruebas elaboradas para los alumnos de 16 años (GCSEs, y previamente a éstas, «Niveles O» y los exámenes CSE), pero el Currículo Nacional lideró el desarrollo de un sistema nacional de medición y de evaluación del profesorado al final de cada etapa clave. Esto significa que actualmente no hay resultados de pruebas anuales de cada cohorte, en su lugar las pruebas miden el rendimiento en periodos que oscilan entre los 2 y los 4 años.

Los logros académicos en las etapas de 1 a 3 están evaluados con niveles nacionales de currículum referidos a criterio. El sistema de medida depende de calificaciones independientes y del Curriculum Authority (QCA) y de la Agencia Nacional de Evaluación (NAA). Los niveles del test están equiparados entre años utilizando anclajes y pruebas de pre-test. Ellos han diseñado lo que en el nivel 4 de la etapa clave 2 es

equivalente al nivel 4 de la etapa clave 3, sin embargo los modelos de valor añadido no dependen de una equiparación vertical o media del crecimiento en una escala común. En la etapa clave 4 hay un rango de calificaciones que pueden compararse utilizando una escala común, pero esta escala no está relacionada con los niveles de las etapas previas. Más información sobre las pruebas puede encontrarse en Ray (2006) y en la página web de QCA.

Es posible el uso de datos del nivel de escuelas, por ejemplo, en la proporción de alumnos que alcanzan un determinado estándar, en los que consideran conjuntamente patrones nacionales y para conocer las diferencias locales entre escuelas. Sin embargo, las bases para los análisis en los últimos años han utilizado datos de tests de los alumnos individuales, donde sus resultados en una determinada etapa clave están vinculados al rendimiento previo en el año anterior. Una vez que se establecieron tests comparables al nivel nacional y el sistema estableció cómo recoger los resultados, pudo ser posible crear bases de datos vinculadas, donde el progreso de un alumno individual o grupo de estudiantes pudo ser medido.

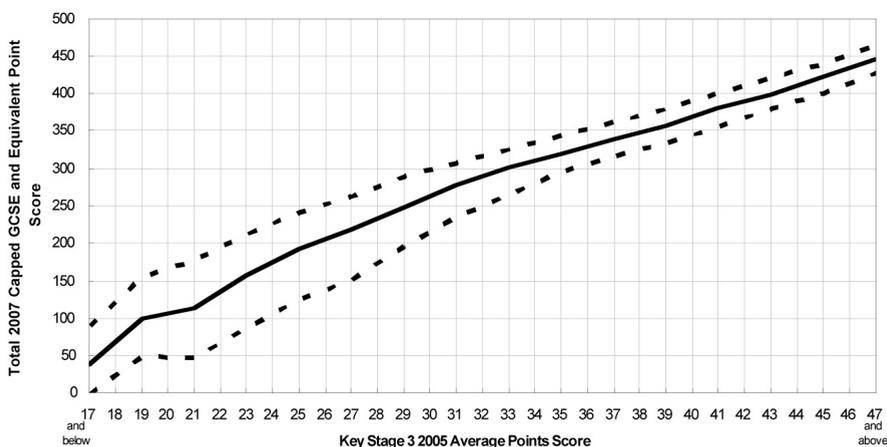
## El desarrollo de los análisis para la mejora de la escuela

Previo a la recolección nacional de datos de tests vinculados, para grupos de escuelas se emplearon análisis de valor añadido para la mejora de las escuelas si sus alumnos realizaban tests específicos, tales como en el programa desarrollado por el centro CEM de la Universidad de Durham (Tymms y Coe, 2003) y en estudios como el de Mortimore et al. (1988) y el de Goldstein et al. (1993). Se puede encontrar un resumen de esos estudios previos en SCAA (School Curriculum and Assessment Authority) (1994). Sin embargo, ninguna de las escuelas tuvo acceso o hizo uso de este tipo de datos. Para solventar este problema, el primer Autumn Package fue realizado en 1998, con patrones nacionales de valor añadido y con patrones estadísticos para grupos de escuelas. Esto permitió a las escuelas comparar sus rendimientos y establecer objetivos<sup>1</sup>. Se desarrolló una versión interactiva del programa, donde las escuelas podían introducir sus datos y ver el abanico de análisis útiles de los alumnos o de los grupos de alumnos.

---

<sup>1</sup> Véase DEPARTMENT FOR CHILDREN SCHOOLS AND FAMILIES, DfEE (1998). La información limitada de valor añadido estuvo disponible un año antes (DfEE/QCA, 1998); antes no era posible vincular los niveles de rendimiento previo de los alumnos con los resultados en las etapas clave de las escuelas.

GRÁFICO I. Ejemplo de gráfico de valor añadido realizada por el Autumn Package



Las tablas y gráficos nacionales del Autumn Package, de los que el Gráfico I es un ejemplo, no requerían una cobertura completa y una precisión en los vínculos. Sin embargo, para aportar datos de valor añadido para las escuelas individuales, fue importante usar correctamente los datos vinculados para todos los alumnos. En 1997, el gobierno estableció el desarrollo de mejores datos del nivel de alumnos y buscó la introducción de un identificador único para cada alumno que podría ayudar a que los datos fueran seguidos a través del sistema educativo. Estos identificadores únicos del alumno fueron introducidos en 1999, siguiendo con las consideraciones prácticas sobre protección de datos. Sin embargo, los datos vinculados fueron completándose de forma creciente, este desarrollo facilitó la centralización del cálculo y aportó indicadores simples de valor añadido para todas las escuelas primarias y secundarias públicas. Estos indicadores fueron desarrollados consultando e introducidos en las tablas anuales de rendimiento comprendidas entre 2002 y 2004. También fueron usadas por los servicios de inspección escolar, Ofsted, y por organizaciones externas encargadas de crear estándares para las escuelas a través de las Estrategias Nacionales de Primaria y Secundaria. Para una exposición más detallada de esta fase inicial de los indicadores de valor añadido se puede consultar más información en Ray (2006).

## La nueva relación con las escuelas

Varios factores llevaron a la siguiente fase de desarrollo de información de valor añadido centralizada. En 2002 se introdujo el Pupil Level Annual School Census (PLASC), ampliando el rango de los datos disponibles sobre las características de los estudiantes de forma centralizada desde considerar sólo género a incluir información útil sobre raza, condiciones sociales, privación, necesidades educativas especiales y primera lengua. Estos nuevos datos eran útiles por sí mismos, pero también ofrecieron la posibilidad de desarrollar modelos de valor añadido que tenían en cuenta esos factores «contextuales». Se estableció un programa de consulta y desarrollo para concebir un nuevo conjunto de «modelos de valor añadido contextuales» que pudieran ser utilizados para aportar información adicional sobre la eficacia de las escuelas. Éstos fueron parte central del New Relationship with Schools lanzado en 2004.

New Relationship with Schools también introdujo School Improvement Partners (SIP's), expertos (a menudo ex directores de centros) quienes podrían trabajar con las escuelas en su autoevaluación. La aproximación tomada por la inspección escolar de Ofsted también varió y se pidió a los inspectores que consideraran un mayor rango de datos de rendimiento, incluyendo el valor añadido. Tuvieron inicialmente que depender de informes escritos (PANDAs) provenientes de un conjunto de estadísticas claves de cada escuela. Ahora, la finalidad era desarrollar PANDAs y el Autumn Package en un nuevo software que sería pilotado con los datos y gráficos de cada escuela y que mostraría un abanico de estadísticas, incluyendo el nuevo valor añadido contextualizado. Este producto, RAISEonline<sup>2</sup> fue introducido en 2006 y se presenta a continuación.

## Los modelos de valor añadido contextual

Una vez que los datos de PLASC se vincularon con la información del progreso del alumno, se pusieron a disposición de la sociedad, hacia finales del año 2002. Estadísticos en el Ministerio de Educación comenzaron a analizarlos para entender las relaciones entre las variables y qué podrían decir sobre el rendimiento nacional. Los planteamientos

---

<sup>2</sup> Puede verse en <https://www.raiseonline.org/>

de una selección de académicos fueron consultados para establecer las futuras direcciones de los trabajos sobre valor añadido y, aunque no se encontró consenso entre sus opiniones, hubo un fuerte apoyo para el desarrollo de modelos más complejos que los utilizados, con los nuevos datos. Fuera del Ministerio de Educación, los estadísticos habían construido modelos de valor añadido que consideraban factores contextuales, por ejemplo el Fischer Family Trust. Asesorados por la Oficina Nacional de Auditorías (2003) se desarrolló información de rendimiento para escuelas que tenían en cuenta no sólo el rendimiento previo, sino también «otras influencias externas del rendimiento», basadas en los datos de PLASC.

En octubre de 2004 se presentó un prototipo contextualizado de modelo de valor añadido (CVA) a las escuelas. Al año siguiente un sistema de puntuaciones CVA fue pilotado para usarlo en las tablas de rendimiento y utilizado por primera vez para la mejora de las escuelas en el año siguiente. El modelo CVA 2005 para escuelas secundarias cubría el progreso de los estudiantes desde el final de la etapa clave 2 hasta la etapa clave 4, y se encuentra desarrollado con detalle en Ray (2006). El modelo equivalente para 2007 se presenta en la Tabla I.

El modelo de regresión de valor añadido más simple (1), sin factores contextuales, incluiría  $y_{ij(t_2)}$  que son los resultados para la etapa clave 4, del alumno  $i$  en la escuela  $j$ , y como variable explicativa  $y_{ij(t_1)}$ , que es el rendimiento previo medio en la etapa clave 2 calculado sobre los cinco años anteriores. Los coeficientes de regresión son  $a$  y  $b$ ;  $e_{ij}$  y  $u_j$  son los valores aleatorios, que son independientes, residuos normalmente distribuidos con media cero [0] y varianza constante. El modelo de coeficientes puede estimarse utilizando los estimadores Mínimos Cuadrados Ordinarios (OLS) y la media de los residuos de los alumnos para cada escuela se utilizaría como medida del valor añadido.

$$y_{ij(t_2)} = a + by_{ij(t_1)} + u_j + e_{ij} \quad (1)$$

En el modelo de valor añadido contextualizado para escuelas sostenidas con fondos públicos se emplean actualmente cuatro variables para crear modelos del modelo de rendimiento previo (que se describe a continuación), un total de diez variables contextuales del nivel del alumno y dos del nivel de las escuelas. Como se ve en (2), por ejemplo,  $x_{5ij}$  representa el primer factor contextual del alumno y  $y_{4j}$  es el primero de los dos predictores del nivel de escuela. (Nótese que las variables explicativas están medidas en 2007 y pueden también ser asignadas al subíndice de la etapa clave 4, aunque la mayoría serían igualmente aplicables en momentos previos, como por ejemplo: género, grupo racial).

$$y_{ij(t_2)} = \beta_0 + \beta_1 y_{1ij(t_1)} + \dots + \beta_4 y_{4ij(t_1)} + \beta_5 x_{5ij} + \dots + \beta_p x_{pij} + \beta_q y_{qj(t_1)} + \beta_s y_{sj(t_1)} + \varepsilon_{ij} \quad (2)$$

Más que usar la aproximación de Mínimos Cuadrados Ordinarios, los indicadores CVA se derivan de un modelo multinivel (Goldstein, 2003), obtenidos con MLwiN, un paquete estadístico que estima los efectos fijos –rendimiento previo y factores contextuales– y calcula los residuos de nivel 1 y 2<sup>3</sup>. Los residuos de nivel 1 muestran la variación de los resultados de los alumnos en relación con sus escuelas. Los residuos de nivel 2 muestran los resultados de las escuelas en relación con los resultados nacionales esperados, dados los factores medidos por los efectos fijos. Estos residuos del nivel 2 son las puntuaciones de valor añadido.

## Factores controlados por el modelo

La variable dependiente «modelizada» es la puntuación total obtenida en la etapa clave 4, basada en las puntuaciones de cada alumno en sus mejores ocho calificaciones. Esto aporta una medida continua del logro desde el alumno más capaz, calificado con ocho A\* o más hasta los alumnos que rinden menos, calificados con las menores puntuaciones (G) e incluye calificaciones de bajo nivel (tales como en alfabetización básica y numeración) tomadas principalmente de alumnos con necesidades educativas especiales.

Las medidas de los resultados fueron elegidas más allá de un simple límite (como el que se establecería si un alumno ha obtenido cinco o más A\* -C), así el modelo resultante refleja bajo –o sobre– rendimiento de los alumnos en todas las partes de la escala más que centrarse sólo en los estudiantes situados en el límite marcado por las puntuaciones C/D. También aporta un útil diagnóstico en el nivel de alumnos, mostrando lo que cada estudiante ha logrado en comparación con alumnos «similares» a nivel nacional. Además de establecer la puntuación hasta las ocho mejores ocho busca capturar tanto la cantidad como la calidad de las puntuaciones en las asignaturas, sin aportar incentivos a las escuelas para admitir alumnos con un excesivo número de calificaciones.

La variable explicativa más importante es el rendimiento previo de cada estudiante y este ha sido «modelizado» de forma tan precisa como ha sido posible. La puntuación

<sup>3</sup> Para más información sobre MLwin, véase: <http://www.mlwin.com/features/index.html>.

media (APS), basada en las notas obtenidas en Inglés, Matemáticas y Ciencias, es usada, junto con otros dos términos adicionales para medir la «diferencia» entre los resultados de Inglés y Matemáticas. Esto es matemáticamente equivalente a incluir cada una de las asignaturas de forma separada, pero tiene la ventaja de que es menos compleja en el manejo de los datos perdidos y también en la simplificación de la presentación. Se incluye también el término cuadrático de APS, reflejando el hecho de que la relación entre los resultados de la etapa clave 4 y la etapa clave 2, anterior, sea no lineal. También realizamos un ajuste posterior para los efectos «techo» y «suelo» para tratar con grupos muy pequeños de estudiantes en los extremos de la escala donde las predicciones del modelo están fuera del rango del rendimiento recogido por la escala de medida. Hacer ajustes en el techo de la escala es importante, aunque sólo el 2% de los alumnos están en ese rango, su tratamiento tiene un efecto desproporcionado en el pequeño número de selectas «grammar schools» que existen en ciertas áreas. Más información está disponible en la página web de Tablas de Logro y de Rendimiento DCSF: <http://www.dcsf.gov.uk/performancetables/>

Las decisiones sobre qué variables contextuales había que añadir en el modelo tras controlar el rendimiento previo estuvieron basadas en un conjunto mixto de consideraciones estadísticas, educativas y prácticas. Dada la necesidad de aportar información sobre el valor añadido para cada escuela es necesario restringir la elección de la información en los datos nacionales. Así, la finalidad fue generar estimaciones de los efectos de la escuela en el aprendizaje desde los residuos del modelo, e incluir las variables explicativas necesarias para recoger los factores que están fuera del control de la escuela. La elección de las variables contextuales del nivel del alumno tomadas en cuenta son las que han mostrado investigaciones internas y externas sobre los factores que explican la varianza de los resultados de una prueba. La Tabla I aporta la relación completa de variables explicativas y sus coeficientes (para ver la Tabla I, váyase al final del artículo).

No se recogen datos de clase social, ingresos familiares o niveles educativos de los padres a nivel nacional de todos los alumnos. Sin embargo, se dispone de alguna información relacionada con niveles de deprivación. Los niños cuyos padres reciben asistencia social y económica, y algunos beneficios relacionados, pueden reclamar ayuda para el comedor (Free School Meals, FSM). Además del FSM, los modelos utilizan algunas otras medidas de deprivación social que proceden del hecho de que los censos escolares recopilan los códigos postales de los domicilios de los alumnos que pueden vincularse a datos del área local. Se han intentado probar varios indicadores locales pero actualmente los modelos utilizan una medida llamada IDACI -Índice de Deprivación Económica que afecta a los Niños (Income Deprivation Affecting Children

Index)-. Es el porcentaje de niños con menos de 16 años en el área local que viven en familias que reciben ciertos beneficios o con bajos ingresos.

Las necesidades educativas especiales cubren un amplio rango de necesidades que a menudo están interrelacionadas también con necesidades específicas que habitualmente se relacionan con particulares tipos de prejuicios. Los niños con necesidades educativas especiales tendrán necesidades y requerimientos que pueden caer al menos en una de estas cuatro áreas: comunicación e interacción; cognición y aprendizaje; desarrollo conductual, emocional y social; y necesidades sensoriales o físicas. El modelo distingue entre dos niveles de necesidades educativas especiales: «Escuela Acción» (donde el profesor de la clase o el coordinador de necesidades educativas especiales aportan ayuda especial) y «Escuela Acción Plus», donde la intervención de la Escuela Acción no ha producido mejoras y se precisa asesoramiento externo.

El censo escolar recoge datos de 18 grupos étnicos, con un código 19 disponible para «no clasificados» dado que la información sobre este dato es voluntaria. Todos estos grupos fueron incluidos en el modelo como indicadores individuales. Una categoría étnica no significativa fue incluida porque faltaba por razones prácticas y de presentación, sería mejor incluir todas las categorías más que combinar dos de ellas con otros grupos sobre la base de los datos de un año (incluir estas variables produce muy pocas diferencias en el modelo general). En definitiva, hay términos de interacción para los grupos étnicos y el estatus de ayuda al comedor.

Las otras variables del nivel de alumnos abarcan un abanico de factores que están asociados con distintas tasas de progreso. Además, en Inglaterra, dado que la mayoría de los alumnos del mismo grupo anual han nacido dentro del mismo año, su mes de nacimiento es significativo: los alumnos que han nacido más tarde tienden a tener resultados más bajos pero realizan progresos más rápidos. Otra variable asociada con bajos resultados pero con rápido progreso debida en parte al dominio de la primera lengua de los alumnos -Inglés y «otra lengua distinta al inglés»-. Ahora incluimos una interacción de términos para el Inglés y un lenguaje adicional (EAL) y los niveles de rendimiento previo. La movilidad de los estudiantes tiende a reducir el progreso y el modelo CVA toma este hecho en cuenta al utilizar el dato del censo escolar sobre la fecha de entrada en la escuela de cada alumno. Finalmente, este es un indicador para los alumnos que están al cuidado de sus autoridades locales, viviendo con padres de acogida localizados en hogares o en otras formas de cuidado residencial o situándose en casas con sus padres.

Hay ventajas y desventajas en la inclusión de variables contextuales del nivel de escuelas. Por ejemplo, se pueden controlar cuestiones adicionales que afectan a todos

los alumnos en la escuela, pero pueden ser parcialmente endógenas, al medir factores que pueden estar influenciados por la propia eficacia de la escuela. Ellos sitúan a las escuelas en un nivel más igualitario de tal manera que las estimaciones de valor añadido pueden aportar un indicador más justo de eficacia escolar, pero al hacer este ajuste, la información puede no ser útil para los padres u otros sectores que están menos interesados en realizar comparaciones sobre esta base. Los modelos CVA se usan para una variedad de finalidades y la decisión fue tomada para incluir dos factores del nivel de escuelas en los modelos para las escuelas secundarias, describiendo la media general del rendimiento previo en el acceso a la secundaria, pero no se incluyen factores adicionales como el nivel de mestizaje étnico en la escuela, el nivel de deprivación social y otros varios más. Estas variables fueron los niveles generales medios y la dispersión (medida a través de la desviación típica) del rendimiento previo APS.

La modelización multinivel (MLM), teniendo en cuenta la estructura de los datos educativos, ofrece un conjunto de complejos modelos más allá de los Mínimos Cuadrados Ordinarios (OLS). La complejidad añadida es vista en ocasiones como una desventaja, particularmente cuando los modelos que resultan son bastante similares, en ambos términos de la estimación de los efectos fijos y de los residuos de valor añadido. En Inglaterra ambas versiones de los modelos OLS y MLM han sido calculadas y comparadas (Ray, 2006). La principal diferencia práctica es que las estimaciones de los modelos de valor añadido multinivel incorporan la concentración (shrinkage). El grado de concentración depende del tamaño de la escuela: las escuelas más pequeñas se concentran hacia la media nacional. No hay una solución fácil para el problema de interpretar el valor añadido en las escuelas pequeñas. La restricción del valor añadido de una escuela a partir de un determinado tamaño significa que la mejora de las escuelas y los sistemas de rendición de cuentas dejan a algunas escuelas fuera. Hacer promedios por más de un año puede aportar figuras más robustas pero impide las comparaciones anuales.

## El modelo para escuelas primarias

El modelo empleado para las escuelas primarias es bastante parecido al modelo de secundaria descrito previamente. La principal diferencia es que los términos del nivel de escuelas se omiten: esto es debido a que el patrón de ingresos es menos variable

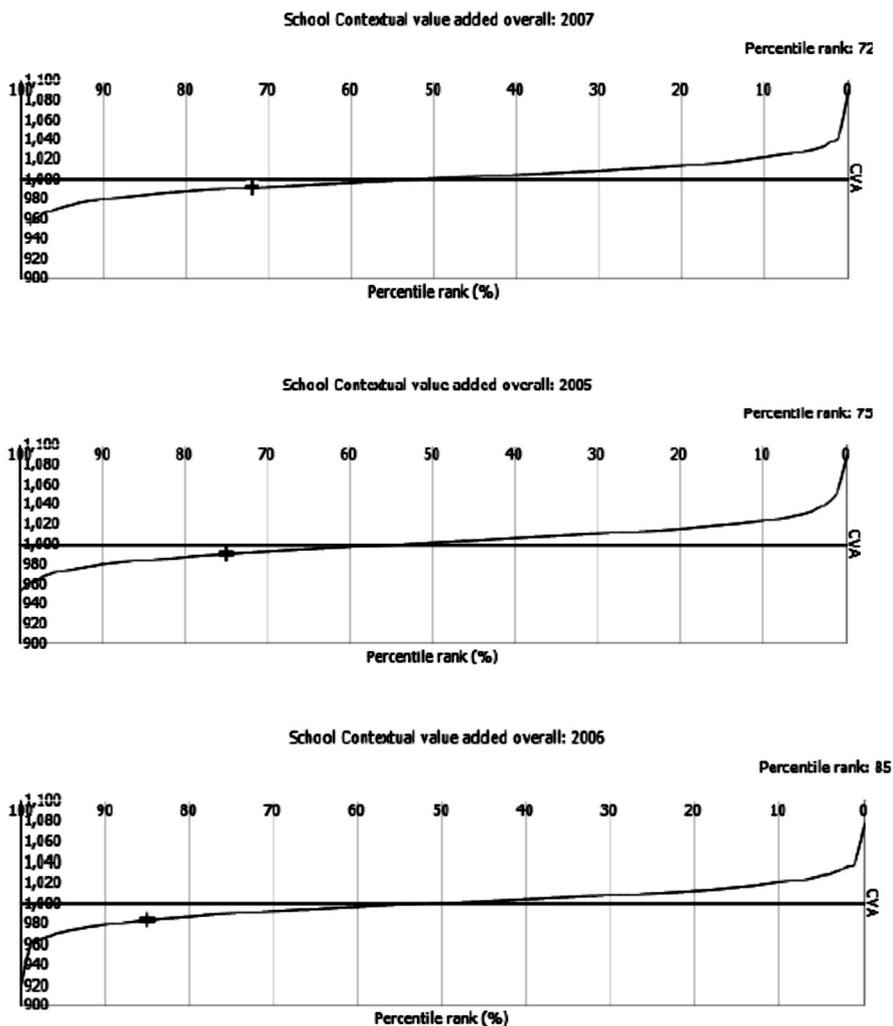
que en las escuelas secundarias, y también porque, con menores números de estudiantes en cada cohorte en primaria, el uso de la misma variable de rendimiento previo en dos niveles puede llevar a modelos más inestables.

## **El uso de valor añadido contextual con RAISEonline**

RAISEonline fue introducido en 2006. Aporta un rango más extenso de datos que las tablas de rendimiento, incluyendo valor añadido para una amplia gama de medidas de resultados y para subgrupos de alumnos dentro de cada escuela. Las escuelas usan RAISEonline como una parte de su autoevaluación y en el proceso de establecer metas que adoptan con la ayuda de los School Improvement Partners. Los datos también están a disposición de los inspectores de Ofsted para su empleo en la evaluación de la extensión de las mejoras escolares o en la capacidad de las mismas para mejorar. Sin embargo, las estadísticas detalladas en RAISEonline no están disponibles para el público en general.

RAISEonline es interactivo y ofrece la posibilidad de investigar cualquier aspecto del rendimiento escolar que sea de particular interés. Además, también provee informes estandarizados que pueden imprimirse y que establecen los indicadores claves en un formato común. Este informe completo para una escuela secundaria puede tener, por ejemplo, aproximadamente 80 páginas. Informes excepcionales, más cortos, también están disponibles para las áreas de mayor interés que aparecen con los puntos fuertes o débiles, basadas en las pruebas estadísticas de significación que comparan las figuras de la escuela con las medias nacionales. El informe completo comienza con alguna información sobre la composición de la escuela, por ejemplo, el porcentaje de alumnos en distintos grupos étnicos y una comparación del rendimiento previo de los alumnos con el patrón nacional (las escuelas con un rendimiento previo relativamente superior se espera que obtengan también mejores resultados). El informe aporta las puntuaciones de valor añadido contextual, primero en una tabla y después en varios gráficos.

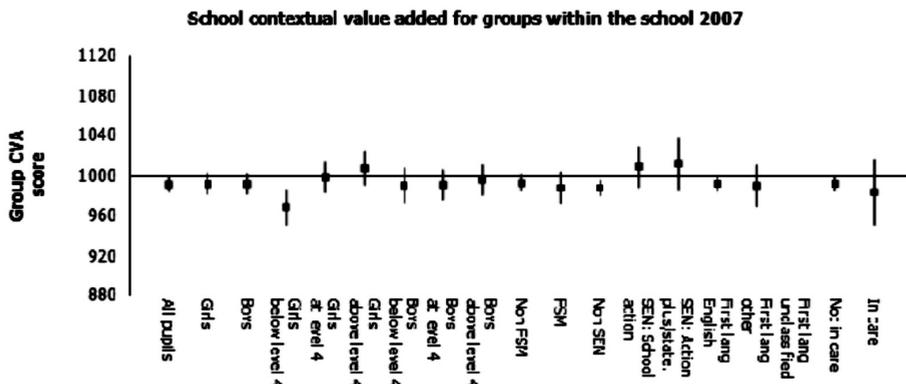
GRÁFICO II. Valor añadido contextual en las etapas clave 2 a 4. General



El Gráfico II reproduce un ejemplo de tres gráficos que se muestran en una página de un informe completo. Los gráficos muestran el valor añadido contextual en diferentes años –gráficos similares están disponibles para distintas asignaturas-. El texto que acompaña indica que: «esta sección aporta la información general de la medida del valor añadido contextual para la escuela respecto a una media nacional de 1.000. La escuela está situada dentro de la distribución nacional para ilustrar el rango de

puntuaciones CVA obtenidas por otras escuelas públicas del conjunto principal. Si el intervalo de confianza para el 95% no se cruza con la línea de la media general, la escuela difiere significativamente de la media general». La CVA para esta escuela aparece por debajo de la media en términos generales, sin embargo, no claramente por debajo en 2006. Una de las cuestiones esenciales para las medidas que están basadas no sólo en una cohorte es si los cambios año a año pueden realmente significar cambios en la eficacia de la escuela. Aquí, por ejemplo, el usuario puede buscar explicación al aparente descenso en 2006 al considerar la naturaleza de la cohorte o de la enseñanza en ese año, o concluir que esta diferencia tan pequeña probablemente no es significativa desde el punto de vista educativo y así centrarse en una variedad de cuestiones sobre por qué la escuela no está por encima de la media en ningún año.

GRÁFICO III. Grupos de alumnos-CVA general

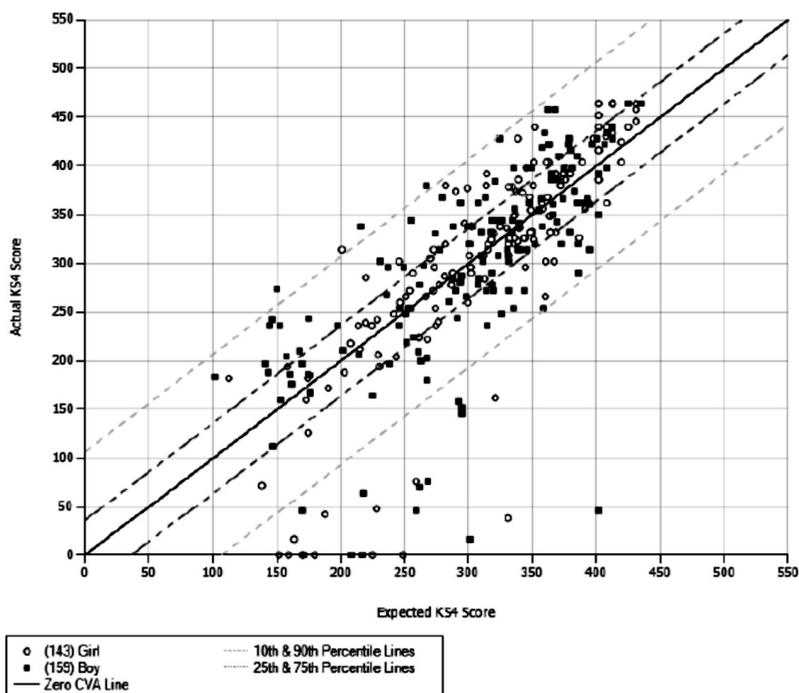


RAISEonline también ofrece valor añadido contextual para subgrupos de alumnos dentro de una escuela. El Gráfico III reproduce un gráfico de ejemplo de un informe completo. Con subgrupos de alumnos es importante reconocer que las pequeñas desviaciones con respecto a la media nacional son difícilmente significativas. Este gráfico muestra que las niñas con bajo rendimiento previo realizan un progreso significativamente menor que lo que se podría esperar (hay 38 niñas en esta categoría).

El Gráfico IV aporta un tercer ejemplo del uso del valor añadido contextual en RAISEonline. Se comparan los actuales resultados de un alumno individual con los resultados que se le predecirían en un modelo de regresión. Las líneas muestran la media nacional y la escuela puede identificar a los alumnos cuyos resultados fueron mayores o menores de lo esperado. Dentro de RAISEonline es posible hacer clic en uno de los

puntos de este gráfico y el sistema ofrecerá detalles de ese alumno en particular. El gráfico también puede filtrar: es posible seleccionar un grupo específico, como los alumnos con ayuda de comedor, y RAISEonline destacará a esos alumnos específicos en el gráfico.

GRÁFICO IV. CVA en etapas clave 2-4. General, predichos frente a los actuales para alumnos (2007)



Coverage 93%

## Otros usos del valor añadido en la mejora escolar

Además del uso en las escuelas, Ofsted y School Improvement Partners hacen RAISEonline. La disponibilidad de las estadísticas de valor añadido aporta una potente fuente de información para que las autoridades locales y el gobierno central apoyen un abanico de actividades e iniciativas. Por ejemplo, ha sido posible identificar aparente bajo rendimiento en las escuelas, utilizando el valor añadido contextual y aportar apoyo orientado

al programa RATL (Raising Achievement Transforming Learning, Transformación de la Enseñanza para el Incremento del Rendimiento). La disponibilidad de los resultados de valor añadido también ofrece información para monitorizar el progreso de grupos de escuelas con respecto a políticas específicas y a acuerdos administrativos.

La publicación anual del valor añadido en las tablas de logro y rendimiento de las escuelas e institutos aporta datos nacionales consistentes sobre el rendimiento de las escuelas, para informar a los padres y al público en general, y asegura que las escuelas sean las responsables de sus resultados. Las tablas son un recurso intensivo producido de forma precisa y anual y mantienen deliberadamente un abanico limitado de indicadores clave. Por tanto, no aportan, por ejemplo, un rango completo de resultados vistos en RAISEonline –los usuarios se dirigen a la inspección Ofsted para tener una imagen completa de una determinada escuela–.

Sin embargo, el valor añadido no está disponible a nivel nacional para estudiar la contribución realizada por los profesores individuales, se puede buscar dentro de las escuelas donde la responsabilidad sobre la enseñanza de materias concretas es clara. El valor añadido contextual puede también contribuir a la mejora de la escuela a través de una vinculación indirecta con el salario, por lo que muchos profesores individuales han progresado hacia una mayor escala de salario a través de la acreditación del progreso. Por este motivo, se les pide que aporten evidencia de que es un resultado de su enseñanza: «sus alumnos rinden bien con respecto a su rendimiento previo, realizando progresos tan buenos o mejores respecto de alumnos similares a nivel nacional».

Los modelos de valor añadido tienen por sí mismos una política de uso, demostrando la relativa importancia de factores que impactan en el progreso de los alumnos, y mostrando que los efectos de la escuela son, de hecho, una parte relativamente pequeña de la variación total. Por consiguiente, abordar los déficits del rendimiento no debe centrarse simplemente en las escuelas con peor rendimiento sino que también necesita considerar el bajo rendimiento «dentro» de las escuelas.

## Conclusión

La disponibilidad de datos nacionales y, desde 2002, de información sobre las características de los alumnos individuales de todas las escuelas públicas, ha permitido que el sistema educativo inglés desarrolle varios indicadores de valor añadido para las escuelas

y grupos de alumnos dentro de las escuelas. Éstos tienen ahora una importancia capital en las políticas escolares de mejora y aportan información importante para los inspectores escolares, para los gobiernos locales y nacional, para los padres y para el público en general. RAISEonline ofrece un medio para analizar estas estadísticas y este artículo ha aportado ejemplos del tipo de análisis que este paquete estadístico puede ahora contribuir a realizar.

La experiencia en Inglaterra ha demostrado que las medidas de valor añadido pueden desarrollarse y aplicarse y que van de alguna manera más allá de las medidas brutas de rendimiento al tratar de identificar la contribución que hacen las escuelas al progreso de sus alumnos. Sin embargo, el sistema actual es nuevo y todavía necesita cierta discusión sobre las mejores combinaciones de indicadores que se deben utilizar en distintas situaciones y para distintos propósitos. Los aspectos de la aplicación en Inglaterra pueden recomendarse a otros países que buscan desarrollar medidas de valor añadido, así como contribuye a valorar la importancia de pilotar y evaluar, de la consulta con las escuelas y con los expertos en educación y de aportar documentación clara y transparente, así como de las herramientas para ayudar a las escuelas, y a otros usuarios a entender las medidas. El modelo óptimo de uso para un país determinado dependerá de las circunstancias locales y de los datos para analizar.

## Referencias bibliográficas

- DEPARTMENT FOR CHILDREN, SCHOOLS AND FAMILIES/ QUALIFICATIONS AND CURRICULUM AUTHORITY, DFEE/QCA (1998). *The Autumn Package.*, London: Department for Education and Employment.
- (2007). Schools and Pupils in England, January 2007 (Final). *Statistical First Release: SFR 30/2007.*
- GOLDSTEIN, H. (2003). *Multilevel Statistical Models* (3<sup>rd</sup> ed.) London: Arnold.
- GOLDSTEIN, H., RASBASH, J., YANG, M., WOODHOUSE, G., PAN, H., NUTTALL, D. & THOMAS, S. (1993). A Multilevel Analysis of School Examination Results. *Oxford Review of Education*, 19, 425-433.
- MORTIMORE, P., SAMMONS, P., STOLL, L., LEWIS, D. & ECOB, R. (1988). *School Matters: The Junior Years.* Wells: Open Books.
- NATIONAL AUDIT OFFICE (2003). *Making a Difference: Performance of maintained secondary schools in England.* London: The Stationery Office.

RAY, A. (2006). *School Value Added Measures in England: A Paper for the OECD Project on the Development of Value-Added Models in Education Systems*. London: Department for Education and Skills.

SCHOOL CURRICULUM AND ASSESSMENT AUTHORITY (SCAA) (1994). *Value Added Performance Indicators for Schools*. London: School Curriculum and Assessment Authority.

TYMMS, P. & COE, R. (2003). Celebration of the Success of Distributed Research with Schools: The CEM Centre, Durham. *British Educational Research Journal*, 29, 639-653.

## Fuentes electrónicas

DCSF ACHIEVEMENT AND ATTAINMENT TABLES WEBSITE. Consultado de:

<http://www.dcsf.gov.uk/performance/tables/>

MLWIN GUIDANCE. Consultado de: <http://www.mlwin.com/features/index.html>

RAISE ONLINE WEBSITE. Consultado de: <https://www.raiseonline.org/>

**Dirección de contacto:** Andrew Ray. Department for Children, Schools and Families. School Analysis and Research Division. Sanctuary Buildings. Great Smith St, London, SW1P 3BT, United Kingdom. E-mail: Andrew.RAY@dcsf.gsi.gov.uk

**TABLA I.** Modelo de regresión en 2007 para las etapas claves 2-4

Dependant variable = capped KS4 point store

-2logL = 6304400

Number of pupils = 563,908

Explanatory factor	Variable	Estimate	Std. Error
<b>Prior attainment</b>	Intercept	162.1	13.1
	KS2 student APS	-5.94	0.24
	KS2 APS (using fine grades) – squared	0.38	0.01
	KS2 English PS deviation	1.4	0.07
	KS2 Maths PS deviation	-0.11	0.07
<b>Deprivation</b>	Does student have FSM?	-22.9	0.33
<b>Deprivation of pupil's local area</b>	Deprivation indicator – IDACI score	-59.51	0.67
<b>Special Educational Needs</b>	Does student have SEN Statement/ Action Plus?	-65.76	0.38
	Does student have SEN - school action?	-34.37	0.31
<b>In care</b>	Has the student ever been in care at this school?	-27.1	1.17
<b>Mobility</b>	Student joined other than Jul/Aug/Sep?	-23.43	0.42
	Student joined within last 2 yrs?	-73.55	0.63
<b>Gender</b>	Is student female?	14.52	0.19
<b>Age</b>	Age within year	-12.94	0.3
<b>Language</b>	Is English not the student's first language?	-8.32	8.25
<b>English as Additional Language * Prior</b>	EAL interacted with prior attainment	4.93	0.68
<b>Attainment</b>	EAL interacted with prior attainment squared	-0.14	0.01
<b>Ethnic group</b>	White Irish	-3.61	1.56
	White Irish Traveller	-64.92	7.66
	White Gypsy/Roma	-54.15	5.25
	White Other	10.52	0.79
	Mixed White/Black Caribbean	-3.08	1.03
	Mixed White/Black African	8.03	2.19
	Mixed White/Asian	10.81	1.38
	Any other Mixed ethnic group	6.03	1.08
	Indian	24.08	0.83
	Pakistani	17.53	0.92
	Bangladeshi	22.91	1.47
	Any other Asian ethnic group	24.76	1.35
	Black Caribbean	13.31	0.9
	Black African	29.19	1.04
	Any other Black ethnic group	11.16	1.62
	Chinese	32.78	1.77
	Any other ethnic group	20.68	1.39
Unclassified ethnic group	-7.42	0.74	
<b>Ethnic group * FSM</b>	White Irish and FSM	2.56	4.04
	White Irish traveller and FSM	11.36	11.98
	White Gypsy/Roma and FSM	25.62	7.79
	White other and FSM	26.53	1.79
	Mixed White/Black Caribbean and FSM	8.46	2.08
	Mixed White/Black African and FSM	3.87	4.46
	Mixed White/Asian and FSM	8.27	3.54

Explanatory factor	Variable	Estimate	Std. Error
	Any other Mixed ethnic group and FSM	11.24	2.48
	Indian and FSM	16.04	1.94
	Pakistani and FSM	17.3	1.24
	Bangladeshi and FSM	19.48	1.89
	Any other Asian ethnic group and FSM	24.71	2.87
	Black Caribbean and FSM	18.2	1.79
	Black African and FSM	20.46	1.65
	Any other Black ethnic group and FSM	8.69	3.19
	Chinese and FSM	34.11	5.12
	Any other ethnic group and FSM	30.83	2.32
	Unclassified ethnic group and FSM	7.14	1.95
<b>Level of school prior attainment</b>	School KS3 APS (using fine grades) for CVA	1.96	0.36
<b>Spread of school prior attainment</b>	School std dev of KS3 APS for CVA	-4.82	0.92
	<b>Random components:</b>	<b>Estimate</b>	<b>Std. Error</b>
	Between school variance	319.42	8.77
	Within school variance	4135.6	7.81
	Variance partition coefficient	0.08	

# No Child Left Behind: retos metodológicos y recomendaciones para la medida del progreso anual adecuado

## No Child Left Behind: methodological challenges and recommendations for measuring adequate yearly progress

Yeow Meng Thum

Michigan State University, College of Education, MI, USA

### Resumen

Este artículo se ocupa de los retos metodológicos de la rendición de cuentas de estudiantes, escuelas, distritos y estados, tal y como se formula en la nueva ley norteamericana, *No Child Left Behind*, y aporta algunas recomendaciones, basadas en investigaciones recientes, para una aproximación viable de la medida de progreso de las escuelas hacia una meta establecida. Concebido como un documento para la planificación, este artículo busca aportar una plataforma analítica que sea lo suficientemente transparente para que la discusión sobre los procedimientos de medida de la rendición de cuentas que puedan separarse del lado más político del actual debate sobre rendición de cuentas. Se presentan las principales justificaciones para 1) emplear puntuaciones de intervalos, 2) utilizar múltiples resultados, 3) estimar las ganancias en valor añadido a partir de datos longitudinales de los alumnos individuales, 4) exigir modelos basados en la agregación, 5) exigir modelos basados en la inferencia y 6) mantener abierta la *caja negra* de un sistema de rendición de cuentas viable. Dentro del mismo marco teórico, se propone una definición de lo que significa para una escuela «conseguir AYP» en el contexto de la NCLB. Se muestra que esta noción de AYP, denominada «AYP-NCLB», se puede *operacionalizar* como una comparación en cualquier punto temporal de la tasa de crecimiento de una escuela con un mínimo de crecimiento exigido para esa escuela si se quiere

que sea «competente» en 2013-2014. El mismo análisis informa de la proporción de los alumnos que son «competentes» en una escuela cada año, que es el principal interés de las aproximaciones de evaluaciones basadas en estándares de referencia.

*Palabras clave:* sistemas de rendición de cuentas, *No Child Left Behind*, modelos de crecimiento, modelos multinivel multivariados, sistemas de rendición de cuentas.

### Abstract

This paper dwells on the methodological challenges for student, school, district, and state accountability as formulated by the new law and provides some recommendations, based on recent research, for a viable approach for measuring progress of schools toward a set target. Conceived as a planning document, this paper aims to provide an analytic platform that will be transparent enough so that the discussion of the procedures for accountability measurement can be better de-coupled from the more contentious policy side of the current school accountability debate. It outlines the principal rationale for (1) employing scale scores, (2) using multiple outcomes, (3) estimating value-added gains from student-level longitudinal performance data, (4) requiring model-based aggregation, (5) requiring model-based inference, and (6) keeping the *black-box* open in a viable accountability system. Within the same framework, it proposes a definition of what it means for a school to 'make AYP' under NCLB. It shows that this notion of AYP, termed 'AYP-NCLB', can be *operationalized* as a comparison at any point in time of a school's growth rate with a minimum growth required of that school if it is expected to be proficient by 2013-14. The same analysis yields the proportion of the students in a school who are 'proficient' each year, the primary interest of standards-referenced approaches to the assessment.

*Key Words:* accountability system, No Child Left Behind, growth models, multilevel multivariate models.

## Introducción

La Ley *No Child Left Behind* de 2001 (NCLB, 2001) representa el papel del gobierno federal mucho más extenso en la educación pública, al caracterizar la nueva ley las provisiones para reforzar la rendición de cuentas del rendimiento académico. Más allá del requerimiento de la medición anual, la ley busca un método para juzgar la eficacia de las escuelas, instaurando una agenda para el progreso último, y estableciendo una

secuencia de las consecuencias específicas para el fallo. Con esta legislación, el gobierno federal parece satisfacerse a sí mismo con el papel de un árbitro de los objetivos de rendimiento y progreso, dejando el establecimiento de la base de la evidencia para los juicios, –incluyendo las materias curriculares y la elección de instrumentos de evaluación–, a los estados. El objetivo inmediato de la NCLB parece ser el establecimiento de un conjunto de procedimientos que ayudarán a vincular las evaluaciones a través del tiempo, a través de los sistemas, y con componentes de evaluación externa tales como los de *National Assessment of Educational Progress* (NAEP) para aportar alguna validación del sistema. Un instrumento común de medida para monitorizar el progreso educativo de los niños de la nación puede ser la clave para producir una divisa común en la evaluación de la productividad, su falta es para muchos el principal impedimento para hacer crecer un esfuerzo nacional coherente orientado a la mejora del debate sobre la educación pública.

Este artículo se ocupa de los retos metodológicos de la rendición de cuentas de estudiantes, escuelas, distritos y estados, tal y como se formula en la nueva ley, y aporta algunas recomendaciones, basadas en investigaciones recientes, para una aproximación viable de la medida de progreso de las escuelas hacia una meta establecida. Queda fuera de la perspectiva de este artículo los problemas relacionados con la alineación con el currículum, los estándares y los test o la elección de test alternativos u otras formas de medir. De forma más concreta, me centraré en las siguientes dos cuestiones esenciales de un esquema de medida de la rendición de cuentas utilizable, como son:

- Definir, medir y monitorizar el progreso del rendimiento de valor añadido e informar de la productividad de unidades múltiples y anidadas (subgrupos de alumnos, escuelas, distritos o estados), y
- Definir el progreso anual adecuado (*adequate yearly progress, AYP*) recogido en la NCLB y diseñar un procedimiento para apreciar y para comparar el progreso de las unidades de rendición de cuentas en términos de AYP.

Concebido como un documento para la planificación, este artículo busca aportar una plataforma analítica que sea lo suficientemente transparente para que la discusión sobre los procedimientos de medida de la rendición de cuentas que puedan separarse del lado más político del actual debate sobre rendición de cuentas. Se presentan las principales justificaciones para 1) emplear puntuaciones de intervalos, 2) utilizar múltiples resultados, 3) estimar las ganancias en valor añadido a partir de datos longitudinales de los alumnos individuales, 4) exigir modelos basados en la

agregación, 5) exigir modelos basados en la inferencia y 6) mantener abierta la *caja negra* de un sistema de rendición de cuentas viable. Dentro del mismo marco teórico, se propone una definición de lo que significa para una escuela «conseguir AYP» en el contexto de la NCLB. Se muestra que esta noción de AYP, denominada «AYP-NCLB», se puede *operacionalizar* como una comparación en cualquier punto temporal de la tasa de crecimiento de una escuela con un mínimo de crecimiento exigido para esa escuela si se quiere que sea «competente» en 2013-2014. El mismo análisis informa de la proporción de los alumnos que son «competentes» en una escuela cada año, que es el principal interés de las aproximaciones de evaluaciones basadas en estándares de referencia. Además, la estrategia analítica propuesta añade directamente cuestiones referidas a 1) la precisión de las decisiones, 2) la elección del punto de partida para realizar evaluaciones y, 3) las llamadas previsiones de «puerto seguro» en la NCLB.

Como líneas maestras y bosquejos que son, por favor entienda que los argumentos de este artículo están necesariamente abreviados. Las propuestas de este artículo se basan sobretodo en el constante estudio e investigación de muchos académicos pero, en aras de la legibilidad, sólo citaré las fuentes primarias en las que se basa este documento y dejaré al lector consultar las referencias allí contenidas (Thum, 2002; Thum, 2003).

## El uso de las puntuaciones de los test

Como en la mayoría de las aplicaciones de la rendición de cuentas, la NCLB cuenta con el uso de evaluaciones estandarizadas de los estudiantes con las que se dirige inmediatamente la atención hacia el consiguiente debate sobre el uso correcto de las puntuaciones de los test. Ha sido largo e inquietante el debate sobre el uso de las puntuaciones procedentes de pruebas estandarizadas para tomar decisiones educativas centradas en las cuestiones sobre la precisión de las puntuaciones de los test, incluso cuando sólo se han usado test válidos y fiables. No obstante, una puntuación de un test puede ser útil si cuidadosamente ponderamos su validez y su precisión. Vean también el reciente *Standards for Educational and Psychological Testing*, 1999.

Mientras que parece claro que ninguna puntuación de un test determinará perfectamente el nivel de rendimiento del estudiante, éste está basado sin embargo en las respuestas de un alumno a una muestra suficientemente amplia de las tareas que definen el dominio. Por tanto, las imágenes que sólo destacan la inherente imprecisión de los resultados

de los test son alarmistas si se pretende que las puntuaciones de los test hagan perfectamente su trabajo. Una puntuación de test, después de todo, es una estimación. Es simplemente una conjetura informada basada, a pesar de todo, en la explícita e imperfecta evidencia del rendimiento. Para cualquier prueba fiable y válida, una medida de cuán (im)precisa puede ser una puntuación se encuentra en su acompañante error típico de medida (*standard error of measurement*, SEM). Sin tener explícitamente en cuenta la imprecisión de las puntuaciones, las inferencias sobre diferencias, tanto si son positivas como negativas, pueden estar sesgadas y las inferencias tenderían a ser demasiado liberales. La cuestión entonces no es si una puntuación de un test puede equivocar la calificación, sino si está sesgada en algún sentido y cuánto, y qué impacto tendrá esta imprecisión sobre las decisiones individuales. Por estas razones, recomiendo procedimientos que tengan en cuenta explícitamente el error típico de medida de las puntuaciones.

## Definiendo y midiendo la productividad del valor añadido

Un buen entendimiento del cambio en los aprendizajes de un alumno es crítico para mejorar la escolaridad pública. La aproximación del valor añadido al problema de la medida del aprendizaje de los alumnos busca situar el cambio dentro del alumno, aislándolo tan bien como sea posible de los muchos factores omnipresentes relacionados con la historia económica y social del estudiante, y de la agrupación con su escuela y su comunidad. Mientras pueden variar las formulaciones específicas de esta idea básica, y dado que el éxito de las mismas no puede garantizarse en ninguna instancia debido a la relativamente exigente técnica y los requerimientos de calidad de los datos (a la vez que necesarios), una aproximación de valor añadido alberga la mayor promesa para responder a la cuestión:

¿Cómo están aprendiendo los niños en nuestras escuelas?

Sin embargo, la intensidad del actual interés en la medida del valor añadido es relativamente nuevo, los métodos para desarrollar los análisis están relativamente bien establecidos en la literatura sobre métodos de investigación. Thum (2002) revisó recientemente la literatura metodológica sobre la medida del cambio y formuló argumentos para apoyar algunas elecciones críticas en la construcción de un sistema estadístico para observar el progreso académico. Algunas de las principales conclusiones son:

## La métrica importa para medir el cambio

En la base de un sistema viable para medir el cambio está una escala de medida de intervalos como métrica deseable para la equiparación. Mientras que el trabajo analítico se viene haciendo en esta escala subyacente, las categorías que reflejan los niveles ordenados de rendimiento pueden usarse para establecer objetivos y para informar de los resultados. La falta de una escala de intervalos en una métrica de equiparación, que es un requisito métrico mínimo para comparaciones válidas del cambio, llevará a imposibles análisis consistentes de rendición de cuentas *-con o sin valor añadido-*. Es la responsabilidad de los productores de test aportar constantemente la necesaria evidencia para sus escalas, haciendo explícito cualquier cambio en los procedimientos, convenciones y asunciones de los modelos que necesariamente son los componentes de la medida estandarizada, así como apoyar el uso adecuado de sus escalas.

## Resultados múltiples ayudan

Las medidas múltiples sirven para replicar nuestras lecturas sobre un constructo de rendimiento, no simplemente como un límite o artimaña para intencionalmente presentar una meta confusa. Cuando se despliegan adecuadamente, los sistemas que emplean múltiples medidas nos ayudan a triangular un constructo de rendimiento más general que entendemos está de forma imperfecta representado por una única media. Además, la redundancia de la información en las medidas múltiples también ayuda a reducir el impacto de los errores de medida. Los análisis multivariados, aquellos que tratan todas las puntuaciones de los test como resultados de forma simultánea, aportarán un conjunto más coherente de resultados cuando sean comparados con los intentos racionales de integrar análisis separados de pruebas individuales de las asignaturas.

## Para medir el cambio, ganancias estimadas

De las distintas aproximaciones para definir el valor añadido, sólo la ganancia en el nivel del alumno aporta un mapa congruente del cambio en el aprendizaje. La puntuación de la ganancia bruta es simplemente una composición lineal de dos medidas positivamente correlacionadas. Basándonos en el ampliamente aceptado modelo de la «puntuación verdadera», podemos demostrar que, si las medidas compuestas son

relativamente precisas, la puntuación de la ganancia bruta tiene una varianza menor que la suma de las varianzas de cada componente de la medida, debido a la correlación entre las puntuaciones verdaderas. La investigación ha demostrado que la fiabilidad de las ganancias dependerá no sólo de la precisión de sus componentes sino también de la distribución de las ganancias de la población, con el resultado de que las puntuaciones de ganancia no son siempre menos precisas que alguno de sus componentes. Por ejemplo, si claramente observamos amplias ganancias que son todas iguales en magnitud, la fiabilidad de las ganancias observadas -una medida normativa de las diferencias en ganancias más allá del ruido del background- es cero.

Y sin embargo, las ganancias brutas pueden no ser inherentemente poco fiables como se pensaba anteriormente, he recomendado que los procedimientos de rendición de cuentas estimen ganancias. Esto se puede conseguir situando todas las puntuaciones de test al mismo nivel como resultados, en lugar de emplear ganancias brutas como punto de partida para el análisis. Es también fácil de demostrar que, dado que este modelo particular de valor añadido emplea al sujeto como su propio control, los factores del nivel individual (como la raza o una beca de ayuda al comedor) que pueden tener un impacto comparable en el rendimiento de los estudiantes en cada medición, ya no predicen las ganancias que se alcanzan. Sin embargo, las ganancias del nivel de aula o de escuela pueden estar correlacionadas con las medidas del aula o de la escuela de esos mismos factores.

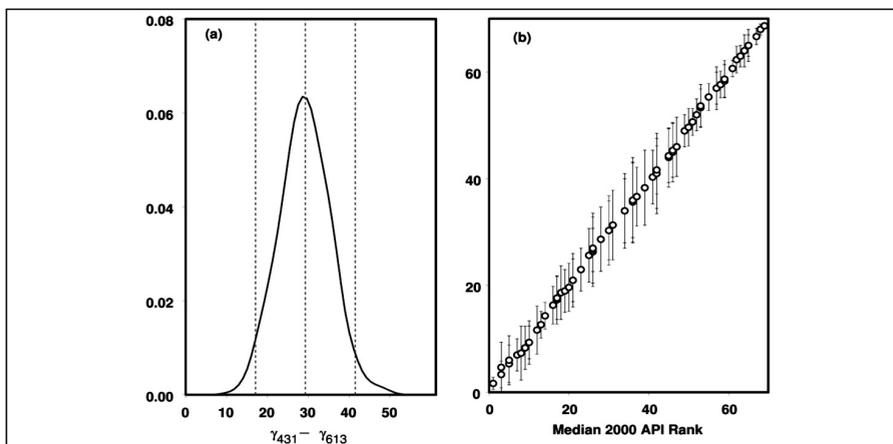
Finalmente, la ganancia estimada no tiene ninguna de las dificultades conceptuales y metodológicas que se esperan de las ganancias residuales obtenidas mediante la regresión de las puntuaciones de los estudiantes del postest sobre el pretest. No hace sólo que sus resultados dependan críticamente del agrupamiento particular de la clase o la escuela, si se usara el pretest como control cuando está correlacionado con el resultado, se violarían las asunciones básicas de la regresión lineal (por ejemplo, que los predictores sean fijos y conocidos, y no correlacionados con los residuos).

## **Exigencia de un modelo basado en agregación**

Los recientes avances en la investigación sobre la eficacia de las escuelas ha demostrado que la historia sobre el nivel de rendimiento del alumno y de la escuela y el progreso cambia, a veces de forma irreconciliable, cuando promediamos las puntuaciones de los test de varias maneras. Es bien conocido que, por ejemplo, la diferencia entre las medias de las puntuaciones del tercer y cuarto curso no siempre son iguales

a la media de las diferencias de las puntuaciones de los alumnos individuales en esos mismos cursos, a menos que el análisis implique a los mismos estudiantes en tercero y cuarto. No sólo la agregación tiene un impacto decisivo en las conclusiones, sino que define la unidad conceptual de qué está siendo medido y, como resultado, esos cambios son los que seguimos. Informar al nivel de la escuela, del distrito o del estado debe realizarse dentro de un marco de modelización coherente a la vez que estadísticamente flexible que comienza con el seguimiento de los cambios en el nivel del estudiante individual y permite una inferencia adecuada de los resultados desagregados.

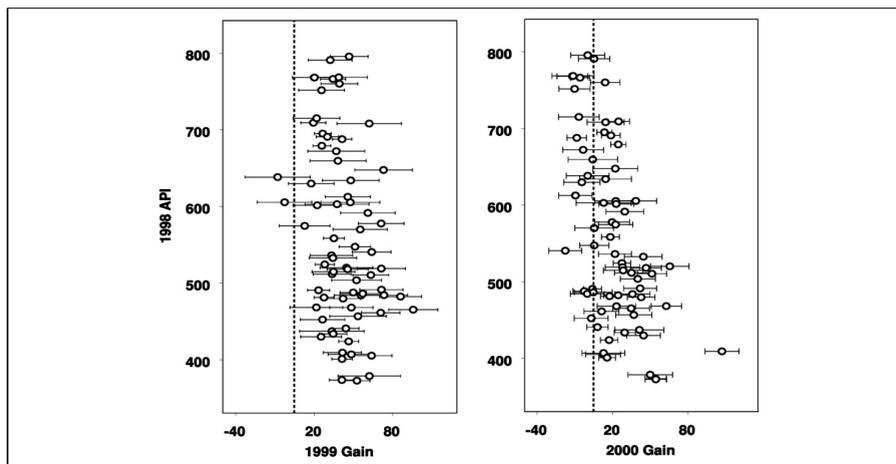
FIGURA I. En California, el API (*Academic Performance Index*) es un índice ponderado compuesto del rendimiento de los estudiantes. a) Comparación de las escuelas 431 y 613 en sus ganancias en API,  $\bar{\gamma}_{431}$  y  $\bar{\gamma}_{613}$ , respectivamente. Las líneas de referencia marcan la diferencia media estimada en el percentil 2,5, en la media y en el percentil 97,5. b) Ordenación (con intervalos de confianza) por la mediana del API en 2000, estableciendo sus intervalos de confianza en el 95%.



### Exigencia de un modelo basado en la inferencia

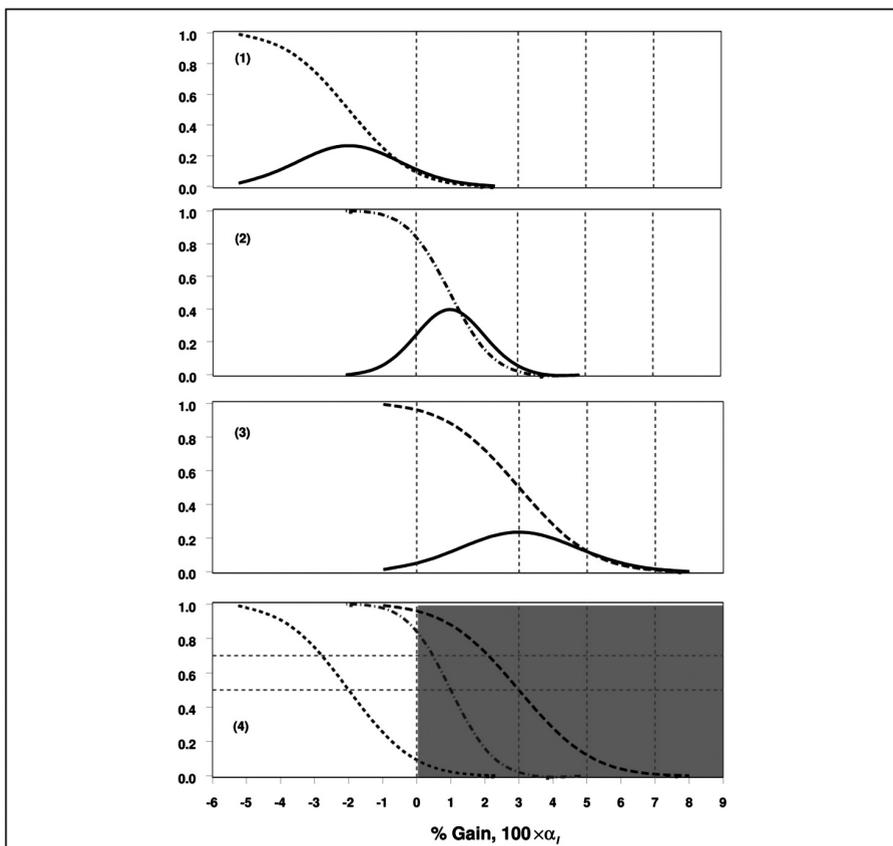
Los sistemas de rendición de cuentas, para ser mínimamente útiles, deben aportar estimaciones defendibles sobre la fiabilidad de sus puntuaciones de productividad. Los cálculos, proyecciones, y ordenaciones del nivel de rendimiento y productividad

**FIGURA II.** Estimaciones de las ganancias del API de las escuelas de LBUSD y las estimaciones del intervalo de confianza del 95% en 1999 y 2000 frente a su nivel en API en 1998. Las ganancias en el API de 1999 correlacionan  $-0,31$  ( $0,12$ ) con el nivel en el API de 1998, mientras que la ganancia del 2000 correlaciona  $-0,52$  ( $0,09$ ) con el nivel de 1998. Escuelas con el mismo nivel en 1998 ganan menos en 2000 ( $-0,10$  puntos) que en 1999 ( $-0,06$ ) en promedio. Las líneas verticales de referencia señalan ganancia 0.



no acompañados por un explícito cómputo de las distintas fuentes de medida y variabilidad muestral deben evitarse porque tales incertidumbres impactarían en las decisiones basadas en las estimaciones brutas. Todas las decisiones de alto impacto deben estar cualificadas por afirmaciones inferenciales construidas adecuadamente para representar completamente la dimensión del uso de la información y el grado de precisión. Estar en disposición de ofrecer estimaciones fiables sobre las puntuaciones de productividad es esencial para un sistema de rendición de cuentas defendible. Las figuras I y II son ejemplos, basadas en una cohorte del Distrito Unificado de Long Beach (LBUSD), de un modelo basado en las comparaciones de las ganancias sobre el valor añadido entre dos escuelas y de los ranking de las estimaciones del nivel de las escuelas, respectivamente. Los perfiles de productividad, tales como los que se muestran en las figuras III y IV, representan una aproximación para presentar como un profesor, una escuela o un distrito están progresando y en que nivel estadístico de confianza, *simultáneamente*, dada la evidencia disponible (Thum, 2002).

**FIGURA III.** Los perfiles de productividad para tres escuelas, en los paneles 1, 2 y 3, están superpuestos en el panel 4 para una comparación más fácil. Cada uno está construido a partir de la distribución marginal posterior simulada de la ganancia de la escuela. Un punto en cada línea indica la ganancia estimada  $100 \times \alpha$  % hecha por una escuela hacia el objetivo de rendimiento (en el eje horizontal) y cómo de fácil es que se obtenga una ganancia observada tan grande como  $100 \times \alpha$  % en términos de probabilidad (eje vertical).



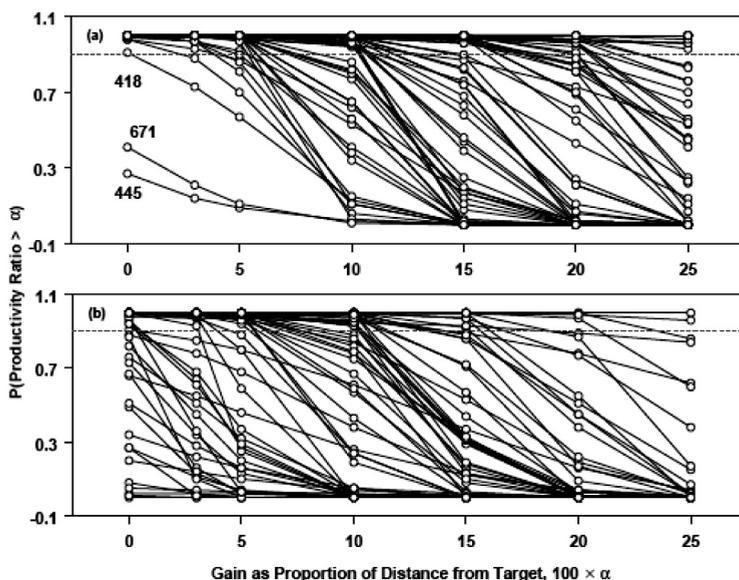
### Mantener abierta la «caja negra»

Parfraseando a la profesora Anita A. Summers de Wharton School, nadie necesita abrir la «caja negra» de un sistema potencialmente útil de rendición de cuentas, del mismo modo que ningún conductor necesita entender cómo funciona un coche para estar

cómodo con su uso, *siempre que algún profesional esté encargado de su diseño y seguridad*. No obstante, necesitamos dejar la llave justo encima de la «caja negra» para facilitar un solícito acceso. Nada debe impedir el desarrollo de técnicas convincentes, en tanto que esté lo suficientemente abierta para la revisión entre pares, la evaluación profesional y la auditoria sistemática de su potencial valor o dañemos el tronco de su desarrollo. Y como la aproximación recibe más y más pruebas realistas en su camino, las inadecuaciones de la metodología o los mitos con respecto a su utilidad práctica, -ambos viejos y nuevos-, serán rápidamente identificados y debidamente superados.

En los párrafos anteriores, he destacado las características de una aproximación profundamente razonada, aproximación que espero sirva para propósitos de diagnóstico y rendición de cuentas, o ambos, mejorando los sistemas que existen actualmente. Estas mismas cualidades la convierten en una fuerte candidata a ser el componente central de un procedimiento de rendición de cuentas. Desgraciadamente, esto también sugiere que la continuada existencia de la mayoría de los sistemas será mucho más difícil de defender desde bases conceptuales y metodológicas.

FIGURA IV. Los perfiles de productividad de LBUUSD reflejados en términos de la PSAA ratio de California, para (a) 1999 y (b) 2000. Las escuelas pueden ahora ser más fácilmente comparadas en términos de su productividad para un nivel de precisión determinado y razonable.



## Progreso anual adecuado

La NCLB exige que todos los estudiantes de las escuelas públicas desde 3° a 8° curso sean competentes en Matemáticas y Comprensión lectora en el año académico 2013-2014, siendo activado el mecanismo de rendición de cuentas en 2005-2006. Esto sugeriría que independientemente del tipo de prueba o del curso específico, el cuerpo de estudiantes estará situado en el nivel de rendimiento de competente en Comprensión lectora y Matemáticas en un tiempo de 12 años. Esto sirve como propósito retórico, pero lo que realmente significa en la práctica es ambiguo. Por ejemplo, no todos los estudiantes cumplirán esta finalidad en el mismo período de tiempo. Y, tal y como la NCLB requiere, si un estudiante comienza la escuela, digamos en 2012-2013, ¿debería ser competente en 2013-2014? ¿Los alumnos que terminaron 8° en 2006 necesitaron ser competentes en ese momento?

Una interpretación razonable podría ser que la NCLB se propone para las escuelas en 2013-2014, no para los alumnos individuales. En el ínterin, son las escuelas las que necesitan mostrar que se han encaminado hacia la competencia en 2013-2014; consecuentemente también necesitan mostrar la importancia de una clara definición de qué significa *lograr una unidad de rendición de cuentas AYP*. La NCLB está primeramente interesada en cómo se mueve cada unidad de rendición de cuentas hacia el objetivo del 100% de competencia, o algún valor que esté aceptablemente próximo a éste, en 2013-2014. La importante cuestión de quién, entre el cuerpo de estudiantes de una escuela, debe ser incluido en la estimación de su productividad en cualquier momento se queda atrás, pero puede ser abordada más tarde.

Muchas de las actuales sugerencias para el AYP giran alrededor de tres ideas (Goertz, 2001). En Texas, las escuelas deben alcanzar un umbral absoluto en rendimiento y en otros criterios. Las metas relativas de crecimiento son empleadas en California. Michigan ofrece un ejemplo en el que el objetivo principal es el descenso de la proporción de estudiantes situados en los niveles más bajos de rendimiento. Todos ellos tienen a su manera un sentido intuitivo y quizá deba ser seguido, privilegiados por algunos estados, al menos en una proporción de sus datos como parte de una estrategia de rendición de cuentas más comprehensiva. Además del intenso interés de muchas agencias estatales, como resultado de la NCLB, está cómo cada una de estas aproximaciones debe ser rearticulada en términos de la NCLB.

No obstante, casi todos los estados establecen una meta de progreso intermedia (mayoritariamente anual) en términos de un nivel de rendimiento o tasa, aunque no todos están claros sobre el marco temporal para lograr este objetivo eventual. Por

ejemplo, California requiere que las escuelas ganen un 5% fijo anual con respecto a su distancia entre su API (*Academic Performance Index*) y el objetivo API del estado fijado en 800, sin ningún requisito temporal para alcanzarlo. La NCLB, por el contrario, establece una línea temporal muy clara, 12 años, para que todos alcancen el nivel competente de rendimiento. Para la NCLB, y por tanto el AYP, debe implicar una solución viable para las cuestiones sobre rendición de cuentas esenciales para las escuelas:

Teniendo en cuenta donde estás en este momento, ¿estás mejorando a un paso que te situará en la meta especificada en el tiempo restante?

Señalaré más adelante cómo una meta explícita y un límite temporal especificado combinan para sugerir una noción del AYP que implica cuestiones de productividad y puntualidad simultáneamente. Específicamente, mi noción de la NCLB sugiere que

El AYP es definido como la tasa mínima de crecimiento basada en la cantidad de bases que una unidad de rendición de cuentas necesita para completarse y alcanzar la competencia en el tiempo restante. En cualquier punto temporal, las unidades de rendición de cuentas logran el AYP si se mejora en una tasa que iguale o exceda este AYP.

He llamado a este nuevo combinado el «AYP-NCLB» para distinguirlo del otro AYP actualmente en uso. En el AYP-NCLB, evalúo el progreso hacia una meta futura con respecto a una línea base de rendimiento relevante. Como mostraré más adelante, nuestro análisis es fácilmente modificable para aportar una evaluación directa sobre si se puede esperar de una escuela si alcanzará el 100% de competencia en 2013-2014.

Además de su claridad conceptual, las tasas generalmente no «saltan» en la manera en que lo hacen las ganancias anuales. También, y lo único de mi aportación, el sistema de rendición de cuentas está preparado para aportar una respuesta directa a la siguiente cuestión planteada por NCLB:

Confiamos al P% de que en este momento su escuela está logrando el AYP-NCLB.

Esta afirmación puede ser claramente transmitida desde el perfil de productividad de la escuela, como se muestra en la Figura III.

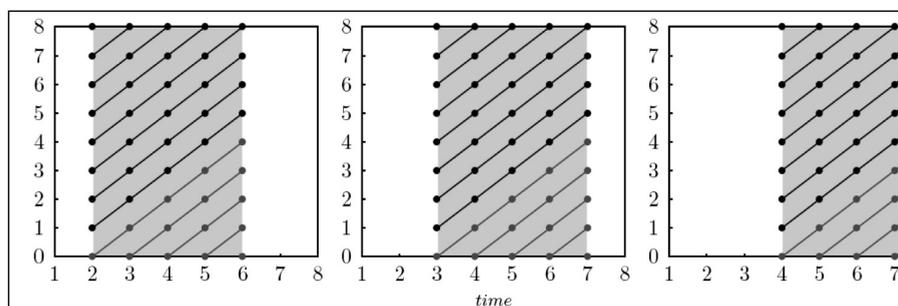
## Los datos son una parte de cualquier definición del AYP

Teniendo en cuenta las decisiones con grandes consecuencias implicadas en el seguimiento de la productividad escolar, necesitamos clarificar todavía más la base de evidencia en la que descansan nuestras estimaciones sobre el rendimiento de las escuelas. Primero, se evita sugerir que la productividad de una escuela es un *rasgo* más que un *estado* (en tanto que un rasgo es considerado una cualidad menos efímera que un estado), necesitamos que el rendimiento de la escuela no esté sólo afectado por los muchos factores relacionados con su composición y recursos, siendo especialmente crítico también reconocer que la productividad de una escuela está circunscrita en el tiempo.

Además, cuando se describe la productividad de una escuela en 2006, por ejemplo, necesitamos un mínimo de claridad tal que nuestras evaluaciones estén basadas en una evidencia disponible entre 2002 y 2006. Otra regla puede ser que se emplee un diseño de bloques móviles (*rolling block*), como el mostrado en la Figura V, si creemos que los datos antiguos pueden no ser relevantes dadas las actuales condiciones de la escuela. Creo que esta práctica es especialmente tentadora porque esperamos que los modelos analíticos y los diseños de datos varíen en los distintos sistemas de rendición de cuentas. De forma específica, sugiero que el bloque de 2002-2006 de los datos de evaluación de los estudiantes de una escuela sean analizados simultáneamente con modelos multivariados mixtos que siga simultáneamente todas las cohortes en cada bloque de datos, cada uno de los cuales es representado por la traza desde abajo a la izquierda hasta arriba a la derecha en la Figura V. Después de determinar que nuestro modelo reproduce adecuadamente los datos, podemos entonces calcular los valores ajustados con las estimaciones de su precisión para todos los nodos, cada materia, curso y año (y subgrupo) específico representado tendrá su estatus, ganancia y tasa de crecimiento. Estos resultados aportan los estadísticos necesarios para personalizar las comparaciones que responden a cuestiones de rendición de cuentas distintas. Por ejemplo, podemos seguir el nivel de rendimiento anual realizado por el tercer curso de una escuela a través del tiempo para tener un sentido del progreso del tercer curso en la escuela. Los cambios en la productividad para el tercer curso en la escuela pueden evaluarse comparando las tasas de crecimiento para distintas cohortes longitudinales cada año. Además, el procedimiento también se acomodará a las comparaciones del crecimiento de la escuela o productividad que tienen en cuenta diversas características de la composición de los estudiantes y de la escuela. Los detalles del modelo para utilizar esta estructura de datos en un sistema de rendición de cuentas como la NCLB se presentan a continuación. Desde mi punto de vista, esta aproximación

es también fácilmente adaptable para sistemas que prefieran el seguimiento de índices de puntuaciones, como el californiano API.

**FIGURA V.** La productividad de una escuela en cada punto temporal debe basarse en un diseño explícito de base de datos de evaluaciones (por ejemplo, 5 años). En eje vertical, el grado 0 comienza en la Escuela Infantil. Nótese que cada traza representa una cohorte de estudiantes longitudinal diferente, los vectores reales de las puntuaciones de los test de los alumnos contendrán valores perdidos.



### **Análisis frente a unidades de rendición de cuentas**

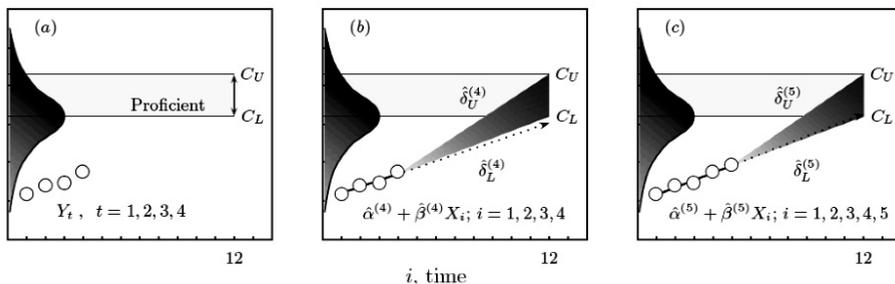
Y, finalmente, necesitamos añadir algunas preocupaciones serias aunque por suerte erróneas desde mi punto de vista, expresadas por muchos de los patrocinadores de los datos longitudinales para la rendición de cuentas. Mientras mi lectura de la NCLB identifica a la escuela y sus subunidades como elementos últimos de la rendición de cuentas, *esto no implica necesariamente que sólo los resultados del nivel de la escuela sean relevantes*. Mi procedimiento propuesto, como señalaba, empleará una base de datos longitudinal de alumnos para caracterizar el AYP-NCLB, en las que *la unidad de análisis sigue siendo (apropiadamente) el estudiante individual, y las unidades de rendición de cuentas son subgrupos de estudiantes y la escuela*.

### **Logrando el AYP bajo la NCLB**

Para medir la productividad en términos del AYP-NCLB, he descrito ampliamente una aproximación para medir el progreso hacia una meta recientemente en Thum (2002) y Thum (2003). Aquí, sólo aportaré un esquema, suprimiendo la notación que sería

necesaria para representar los múltiples resultados y las múltiples unidades de análisis completamente. Para el propósito limitado de este artículo, la lógica esencial está bosquejada con la ayuda de los paneles (a), (b) y (c) de la Figura VI, sólo para una escuela en una única prueba y para un único curso. Se presenta a continuación el tratamiento completo para múltiples medidas y múltiples criterios, incluyendo la sutileza relatada de la forma de la función de predicción y del heterogéneo error de varianza dentro de un modelo para datos anidados.

FIGURA VI. En cualquier punto temporal,  $t$ , la productividad de una escuela es medida simultáneamente junto con su AYP. Los logros en escala de intervalos están dibujados en el eje vertical frente al tiempo en el eje horizontal (ver el texto para mayores explicaciones).



### Puntuaciones y metas

La distribución de puntuaciones para la población se muestra en el fondo de cada gráfica de la figura VI. Las líneas de referencia horizontal marcan las puntuaciones de corte superior e inferior para la meta de rendimiento de «competente», denotados por  $C_L$  y  $C_U$  respectivamente. Los análisis se realizarán utilizando puntuaciones de intervalos, incluso cuando *algunas lecturas erróneas de la legislación puedan sugerir que los análisis implicados necesiten comenzar y terminar con categorías de rendimiento*. Ya que utilizamos estándares de rendimiento que definen las puntuaciones de logro original en intervalos, aportamos una evaluación más consistente internamente de lo que significa, por ejemplo, ser «competente» y, como resultado, se destaca la relevancia directa de cómo se han establecido los estándares de rendimiento. Se recomienda también una alternativa atractiva más consistente internamente cuando se comparan los estándares de rendimiento que están impuestos internamente, tal y como fija un crecimiento anual fijo del 5% en el API de California, que no están vinculados tan próximamente a los dominios que están siendo evaluados.

## Rendimiento y productividad

En el panel (a) de la Figura VI, llamamos «o» al progreso escolar a un resultado de una escuela ( $Y_1, Y_2, Y_3, Y_4$ ) para los primeros 4 años. El límite temporal, de acuerdo con la NCLB, es  $t = 12$ . Utilizando un modelo lineal simple, por ejemplo, podremos aproximar cómo está rindiendo la escuela en el momento  $t$ , estimando  $\hat{Y}_t$  a partir de

$$\hat{\alpha}^{(t)} + \hat{\beta}^{(t)} X_i .$$

Si el predictor  $X_i$  cifrado en el tiempo de tal manera que  $X_i = i - t$  entonces  $\hat{\alpha}^{(t)}$  da una estimación directa de  $\hat{Y}_t$ . Nótese que la tasa de crecimiento,  $\hat{\beta}^{(t)}$ , es simplemente una ganancia media estimada, que en tanto que medida de rendimiento, es menos susceptible a fuertes fluctuaciones comúnmente observadas en las ganancias de año en año. Cuando examinamos la conducta de  $\hat{\beta}^{(t)}$  a través del tiempo, estamos estudiando la productividad de la escuela para un marco relevante de tiempo.

## Definición del AYP-NCLB

Considerando el tiempo restante  $12-t$ , la escuela necesitará crecer en una tasa igual a

$$\hat{\delta}_L^{(t)} = \frac{C_L - \hat{Y}_t}{12 - t}$$

basada en nuestra mejor lectura de dónde está la escuela en el tiempo  $t$  para alcanzar la meta, limitada por abajo por el punto de corte  $C_L$  en  $t = 12$ . De forma similar,

$$\hat{\delta}_U^{(t)} = \frac{C_U - \hat{Y}_t}{12 - t}$$

aporta la tasa necesaria para superar la competencia.  $\hat{\delta}_L^{(t)}$  y  $\hat{\delta}_U^{(t)}$  conforman los límites inferior y superior respectivamente de la «raya» de nuestra mejor estimación del nivel actual de rendimiento de la escuela con respecto a la meta temporal situada en  $t = 12$ . El Panel (b) y el Panel (c) describen las tasas de crecimiento de la escuela y el AYP-NCLB en  $t = 4$  y  $t = 5$ , respectivamente. Una implicación inmediata para casi todas las formulaciones existentes del AYP es que, dado que la meta última es fija en la NCLB, la meta intermedia que es

El AYP-NCLB cambia con el tiempo.

En tanto que seamos capaces de determinar mejor y mejor el progreso de la escuela a través del tiempo, la práctica común de establecer *una tasa anual fija para un período de tiempo extenso tendrá menos sentido*. En definitiva, para ser específico en cada punto temporal, mi razonamiento más bien sugiere que el AYP-NCLB es

una unidad específica de rendición de cuentas y también es específica para un test. La agregación de test y de condiciones subagrupadas aportará la estimación apropiada para evaluar cómo varias unidades de rendición de cuentas satisfacen la NCLB.

### Alcanzando el AYP-NCLB

A pesar de que podamos calcular en cualquier momento la proporción de estudiantes que alcanzan la competencia, siento que esta imagen sea poco representativa de la tendencia de crecimiento escolar y sugiero en su lugar la siguiente alternativa. Dada nuestra mejor estimación de productividad en el momento  $t$ , por ejemplo  $\hat{\beta}^{(t)}$ , la escuela alcanzará el AYP-NCLB si

$$\hat{\beta}^{(t)} \geq \delta_L^{(t)} .$$

Entonces, si una escuela alcanza el AYP-NCLB implica una comparación de tasas de crecimiento. Esta comparación aporta una estimación del eventual rendimiento de la escuela. Esto es, esperamos que la escuela rinda en el nivel de competente en 2013-2014 si

$$\delta_U^{(t)} \geq \hat{\beta}^{(t)} \geq \delta_L^{(t)} .$$

En el Panel (b) y el Panel (c), las proyecciones para alcanzar eventualmente la meta están dadas por las líneas negras con puntos. En nuestro ejemplo, en  $t = 4$ , la escuela parece no alcanzar el AYP-NCLB. En  $t = 5$ , la escuela parece alcanzar el AYP-NCLB. Tener a la escuela retrasada otra vez en  $t = 5$ , podría ser objeto de alguna forma de intervención. Parece claro que, en el AYP-NCLB, continuamente evaluamos el progreso hacia una meta futura de rendimiento fija con respecto a una línea base de rendimiento relevante. Debe estar también claro que comparaciones en múltiples metas, como por ejemplo las definidas para otros niveles de rendimiento o para subgrupos específicos de niveles de rendimiento, son también fácilmente desarrollables.

### Confianza en la decisión

Nuestros resultados lejos de confiar en la comparación de tasas de crecimiento estimada, representan, por un lado, cómo está rindiendo la escuela,  $\hat{\beta}^{(t)}$ , y por otro sirve como punto de referencia interino del AYP-NCLB,  $\delta_L^{(t)}$ . Dado que ambos son estimaciones, es importante, como he señalado detalladamente, para caracterizar el nivel de certeza vinculado a sus comparaciones después de tener en cuenta todas las fuentes de variación conocidas. Esto es especialmente claro en nuestro ejemplo para  $t = 5$ , donde los resultados parecen –para algunos decisores al menos– demasiado próximos para llamar la atención.

De forma más específica, deseáramos conocer cuan plausible es que una escuela logre el AYP-NCLB en  $t$ , o

$$\text{Prob}(\hat{\beta}^{(t)} \geq \hat{\delta}_L^{(t)} | \mathbf{Y}).$$

De forma similar,

$$\text{Prob}(\hat{\delta}_U^{(t)} \geq \hat{\beta}^{(t)} \geq \hat{\delta}_L^{(t)} | \mathbf{Y})$$

aporta la probabilidad de que la escuela sea competente cuando llegue la fecha límite en  $t = 12$ . Al acordarse que las altas demandas están vinculadas a nuestras decisiones de rendición de cuentas y dado un razonable consenso sobre el grado de seguridad que tenemos antes de realizar afirmaciones sobre si una escuela alcanza o no el AYP-NCLB, podemos seleccionar un rango de niveles de confianza, por ejemplo 70%, 80% o 90% para ayudarnos a llegar a una decisión basada en los datos. Si deseáramos tener mayor certeza sobre si una escuela lograra el criterio AYP-NCLB en el momento  $t$ , podríamos seleccionar un nivel de confianza del 90%, entonces la escuela alcanzará el AYP-NCLB si

$$\text{Prob}(\hat{\beta}^{(t)} \geq \hat{\delta}_L^{(t)} | \mathbf{Y}) \geq 0.90 .$$

Si nos sentimos todavía incómodos en la selección de un nivel de confianza particular, podríamos emplear el perfil de productividad descrito previamente en la Figura III para ayudarnos a conducir una afirmación más precisa sobre la confianza estadística de nuestra decisión para un abanico deseable de niveles de confianza.

## Evaluando los programas de reconocimiento

Las decisiones de rendición de cuentas son difíciles, y por tanto, los análisis cuidadosos de datos son críticos, precisamente porque el verdadero nivel de mejora es desconocido. Sin embargo, cuando las estimaciones de productividad escolares están avaladas por estimaciones fiables en la manera descrita previamente, apoyan firmes acciones de rendición de cuentas (premios o sanciones) en el nivel de escuelas. Por ejemplo, si se otorga un premio a una escuela, tendremos al menos el 90% de certidumbre de que la escuela logra o excede su meta de crecimiento, habiendo aportado una base estadística sólida para cómo son reconocidas las escuelas. De forma similar, se puede realizar para subgrupos específicos, o para cualquier combinación arbitraria de subgrupos. Es interesante destacar que este procedimiento también puede servir

como *estándar de oro*, en tanto que no haya otro disponible, de los varios regímenes de reconocimiento. Como ejemplo obvio pero cautivador, podemos entonces evaluar la precisión del examen de un programa de premios alternativo que no tenga en cuenta la precisión y la estimación de la productividad de una escuela si está basado en la comparación del rendimiento escolar con una meta explícita, usando bien una media o una estimación de un porcentaje de corte (PAC). Por último, sugiero, sin desarrollarlo, que esta aproximación para establecer y estimar la precisión de una decisión de rendición de cuentas rodea la especulación directa con respecto al número mágico del tamaño del grupo, el llamado «tamaño mínimo del grupo», exigido para tomar decisiones «estadísticamente fiables».

### Puntos de inicio alternativos

Se han propuesto varias alternativas en la legislación para su uso como líneas de referencia. Por ejemplo, utilizar los datos de 2001-2002 puede considerarse como emplear el nivel predicho de rendimiento en el momento  $t$  del subgrupo con más bajo rendimiento dentro de una escuela, o la escuela con rendimiento más bajo de un distrito, o el nivel medio predicho de un subgrupo de escuelas en el sistema en lugar de  $\hat{Y}_L$  para la definición de  $\hat{\delta}_L^{(t)}$ . De hecho, la comparación de cada escuela con líneas de base múltiples, algunas de las cuales puede ser importantes para algunos subgrupos de alumnos, no plantea cargas adicionales para nuestros análisis.

### EIAPY-NCLB y el % de competentes

Aún cuando he designado en el AYP-NCLB una herramienta de pronóstico en términos del rendimiento medio esperado de los alumnos de una escuela en un momento temporal, se puede fácilmente *informar* de los análisis en términos del nivel de rendimiento estimado para un alumno individual. Supongamos que  $\hat{Y}_i$  denomina a la puntuación predicha para el estudiante  $i$  de una escuela en el momento  $t$ . Entonces el estudiante  $i$  tiene una probabilidad

$$\text{Prob}(\hat{Y}_i \geq C_L | \mathbf{Y})$$

de ser al menos «competente» y el porcentaje estimado de alumnos en una escuela en el momento  $t$  que son al menos «competentes» es simplemente

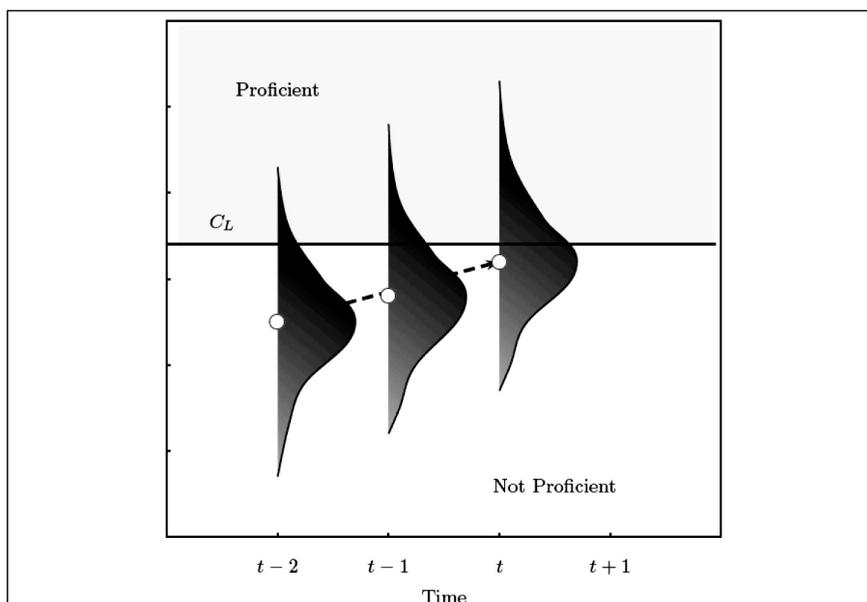
$$\hat{P}_\mu = \frac{100}{n_t} \sum_{i=1}^{n_t} \text{Prob}(\hat{Y}_{it} \geq C_L | \mathbf{Y}),$$

donde  $n_t$  es el número de estudiantes en la escuela en el momento  $t$  (véase la Figura VII). Al utilizar la distribución del rendimiento de la escuela en dos años consecutivos,  $t$  y  $t-1$ , podemos añadir directamente la llamada provisión de «puerto seguro» estimando

$$\text{Prob}(\hat{P}_{jt} - \hat{P}_{j,t-1} \geq 10\% | \mathbf{Y}),$$

para evaluar qué verosimilitud puede tener que haya un decremento de al menos un 10% de alumnos no competentes en la escuela en un periodo de dos años. No hace falta decir que las conclusiones basadas en el AYP-NCLB y los diversos criterios del porcentaje de alumnos pueden divergir en algunos casos. Creo que el último es más conservador y duro para estimar con niveles razonables de precisión.

**FIGURA VII.** Dadas nuestras estimaciones de crecimiento escolar en cualquier punto temporal, se puede seguir simultáneamente la reducción en el porcentaje de estudiantes no competentes de una escuela. Para cada año, la proporción de alumnos no competentes es la proporción de la distribución por debajo de  $C_L$ .



Puesto que la mayoría del lenguaje empleado en esta legislación para describir el progreso está en términos de esta distribución, muchas agencias sienten comprensiblemente la responsabilidad de ofrecer este resumen como punto inicial y final de su sistema de rendición de cuentas. Ya sabemos que cuando empleamos categorías de rendimiento que se obtienen por la politomización de las escalas originales de puntuaciones, estamos despreciando información importante, informando de menores, aunque significativos, cambios en los aprendizajes que permanecerán indetectables. Como es también evidente ahora, los análisis adecuados pueden comenzar con puntuaciones de intervalos longitudinales de los estudiantes y los aspectos de los resultados pueden presentarse en términos de la distribución del rendimiento de la escuela. Comenzando con las distribuciones transversales observadas del rendimiento de la escuela tendríamos extremadamente mal representado el carácter longitudinal del crecimiento y del cambio en los aprendizajes de los alumnos.

## Resumen y conclusiones

Tanto si está establecido explícitamente o no, es importante reconocer que cada análisis implica un modelo y debemos escudriñar con cuidado cualquier análisis que parezca libre de uno. En este artículo, he aportado un boceto de un procedimiento de rendición de cuentas que responde a la NCLB. El sistema propuesto necesita para operar correctamente comenzar con datos longitudinales del logro de los alumnos. Esto requiere que una prueba basada en estándares con logro expresado en una escala de intervalos no sea una restricción real, en ausencia de alternativas reales. Los que desarrollan las pruebas, que deben jugar un papel más activo para apoyar el renovado esfuerzo sobre la rendición de cuentas en toda la nación, deben aportar garantías periódicas de que las puntuaciones de sus test están equiparadas por cursos y a través del tiempo. Deben también ayudar a clarificar los estándares de rendimiento para sus pruebas, por ejemplo, ¿cuáles son las correlaciones del rendimiento con el nivel de «competencia» en sus test? Debe dirigirse un mayor asesoramiento hacia la tutela del uso adecuado de las puntuaciones de sus test. Con test juiciosos medidos en una escala deseable, los elementos de los testados modelos multinivel multivariados pueden formar la esencia analítica para la estimación del crecimiento del rendimiento, para lo que he superpuesto una nueva formulación de procedimientos para hacer inferencias razonables sobre si una escuela alcanza o no el AYP. Concluyo brevemente añadiendo tres preocupaciones ampliamente expresadas:

- Mientras la NCLB configure sus objetivos de logro en términos de categorías de rendimiento, mantengo que esto simplemente facilitará la comunicación siempre que no altere nuestro foco analítico de los resultados continuos de los estudiantes. Por las razones previamente señaladas, tengo serias dudas acerca de las bases conceptuales de las recientes recomendaciones llamadas *evaluaciones referidas a estándares* frente a aproximaciones más «tradicionales» (por ejemplo en Schwarz et al., 2001). Por qué trabajar con información rebajada a categorías de rendimiento cuando tenemos múltiples lecturas del rendimiento de los estudiantes que están sintetizadas por una medida continua de intervalos. No se debe confundir los principios básicos de análisis con la simple conveniencia del informe para construir un sistema de rendición de cuentas sólido. El rendimiento en un criterio y el rendimiento relativo a un conjunto de normas son informaciones complementarias que son igualmente importantes para la salud de la educación pública.
- Necesitamos también evitar agregaciones simplistas de estudiantes inferidas al nivel de escuela porque, sin un modelo explícito de la naturaleza anidada de los datos, las informaciones importantes referidas a los test y las fuentes de variación muestrales se perderán sin posibilidad de recuperación. Y como ya he señalado antes, esto es bastante contrario a muchas de las actuales lecturas de la NCLB: sólo porque la NCLB establece explícitamente metas para las escuelas esto no significa que los resultados de las escuelas deban comenzar por estos análisis. Recomiendo estimaciones del nivel de escuelas sintetizadas a partir de datos longitudinales de los estudiantes y hacer entonces inferencias en función de estos análisis.
- Los estándares altos están bien y son buenos, pero los análisis preliminares en todo el país cuestionan si los objetivos de la NCLB son alcanzables en la práctica. Y esto no quiere decir que debemos establecer objetivos que juzguemos como alcanzables, ya que se puede fácilmente cuestionar su credibilidad. Nótese que, como un estándar que es, el AYP-NCLB establece sólo una senda para buscar un estándar de rendimiento en una determinada cantidad de tiempo. Y mientras podamos tener el impulso para altos estándares por su valor motivacional, y aún cuando el AYP-NCLB aporta un estándar útil para el progreso desde nuestro punto de vista, ciertamente es más necesario conocer el nivel de ganancias de aprendizaje que es posible en muchos contextos de evaluación para llegar a estándares razonables de crecimiento. En este momento, coincido con que las ganancias normativas, como nivel sucesivo de norma, son obviamente piezas que aún están ausentes en el puzzle de la rendición de cuentas.

## Referencias bibliográficas

- AMERICAN EDUCATIONAL RESEARCH ASSOCIATION (AERA), AMERICAN PSYCHOLOGICAL ASSOCIATION (APA), & NATIONAL COUNCIL ON MEASUREMENT IN EDUCATION (NCME) (1999). *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association.
- GOERTZ, M. E. & DUFFY, M., with LE FLOCH (2001). *Assessment and Accountability Systems in the 50 States: 1999-2000*. A CPRE Report.
- No Child Left Behind Act of 2001, Pub. L. No. 107-110, 115 Stat. 1425 (2002).
- SCHWARZ, R. D., YEN, W. M., & SCHAFFER, W. D. (2001). The challenge and attainability of goals for adequate yearly progress. *Educational Measurement: Issues and Practice*, 20, 26-33.
- THUM, Y. M. (2002). Measuring Student and School Progress with the California API. *CSE Technical Report 578*. Los Angeles: Center for Research on Evaluation, Standards, and Student Testing, UCLA.
- THUM, Y. M. (2003). Measuring Progress towards a Goal: Estimating Teacher Productivity using a Multivariate Multilevel Model for Value-Added Analysis. *Sociological Methods & Research*, 32, 153-207.

**Dirección de contacto:** Yeow Meng Thum. Measurement and Quantitative Methods, Counseling Educational Psychology and Special Education, Erickson Hall, Michigan State University, College of Education, East Lansing, 48824 Michigan. E-mail: thum@msu.edu

# Consideraciones de validez prioritaria para la evaluación<sup>1</sup> formativa y de rendición de cuentas

## Priority validity considerations for formative and accountability assessment

Eva L. Baker

Universidad de California en Los Ángeles. National Center for Research on Evaluation, Standards, and Student Testing (CREST). California, Estados Unidos

### Resumen

Este artículo se centra en la validez de los sistemas de evaluación (Baker, 2007a) usados en las escuelas con la intención de mejorar la instrucción, bien a través de evaluaciones formativas o de evaluaciones para la rendición de cuentas. El argumento que se presenta apoya la necesidad de estudios de la sensibilidad instructiva de los tests, necesitando medidas mejores y más equiparadas con la enseñanza en el aula. La agenda de investigación caracteriza a estas preocupaciones como extensas y esenciales. A través de la presentación de los trabajos que se realizan en el CREST (*Center for Research on Evaluation, Standards and Student Testing*), de la Universidad de California en los Ángeles, su directora, Eva Baker, expone la importancia de la sensibilidad instructiva y los aspectos que se derivan de ella para, posteriormente, pasar a exponer una serie de cuestiones relativas al diseño de pruebas referidas a criterio (CRT). Finalmente, se comenta la situación en la actualidad en los Estados Unidos, donde se realizan evaluaciones con base normativa. Planteadas estas cuestiones, se pone en relación dos importantes aspectos relacionados con la evaluación de los aprendizajes, a saber, la sensibilidad

---

<sup>01</sup> El trabajo presentado en este artículo ha sido apoyado por el *National Research and Development Centers, PR/Award Number R305A050004*, y administrado por el *Department of Education's Institute of Education Sciences (IES)* de los EE.UU. Los hallazgos y opiniones expresados en este trabajo no reflejan necesariamente las posiciones o políticas del *National Research and Development Centers* o del *Department of Education's Institute of Education Sciences (IES)* de los EE.UU.

instructiva y las pruebas que persiguen objetivos múltiples. Como conclusión del artículo presentado se presenta un resumen, de modo provisional, sobre las cuestiones que se asumen para la sensibilidad de la instrucción, así como argumentos a favor y en contra de la misma. Se concluye presentando las experiencias del CRESST en el tema.

*Palabras clave:* evaluación y rendición de cuentas, estudios de validez, evaluación formativa, validez instructiva, evaluación educativa.

### **Abstract**

This paper focuses on validity of testing systems (Baker, 2007a) used in schools intended to improve instruction, either through formative or accountability-based assessments. The argument put forward supports the need for studies of the instructional sensitivity of tests to assure that they are actually measuring school effects. The research agenda suggested by these concerns is extensive and essential. Eva L. Baker presents CRESST (Center for Research on Evaluation, Standards and Student Testing, University of California, Los Angeles) research jobs related to instructional validity and main issues related to Criterion Referenced Test (CRT). Finally, in order to conduct such studies, better and more scalable measures of classroom instruction are needed.

*Key Words:* assessment and accountability, validity studies, formative assessment, instructional validity, educational assessment.

## **Introducción**

La rendición de cuentas es una función deseada en todo el mundo, y ha adquirido enorme importancia en el área de política educativa. La rendición de cuentas significa hacer responsable a un individuo, grupo, agencia o gobierno de sus acciones y de sus consecuencias conforme intenta alcanzar unas metas específicas. En los Estados Unidos, la rendición de cuentas en educación (Baker, en prensa; Baker *et al*, 2002) se ha centrado fundamentalmente en las escuelas, especialmente en el rendimiento de los estudiantes en ciertos dominios curriculares (matemáticas, lectura y escritura y ciencias) en los niveles de escuela elemental y secundaria. En otros países, la rendición de cuentas puede enfocarse en un rango más amplio de contenidos y sobre el rendimiento individual. En algunos contextos la rendición de cuentas y las consecuencias del rendimiento inadecuado de alumnos o escuelas se distribuyen ampliamente, y las

sanciones rodean a estudiantes, profesores, administradores, y responsables políticos. Este artículo se centrará en la experiencia norteamericana y ofrecerá la perspectiva desde la investigación y el desarrollo.

Una revisión del progreso de los estudiantes de Estados Unidos, los más recientes impulsos por la evaluación del rendimiento comenzaron alrededor de 1992, no es halagüeña. Habitualmente, los estudiantes norteamericanos rinden cerca de la media en las comparaciones internacionales (OCDE, 2005; PISA, 2003). Más seria es, por supuesto, la brecha de rendimiento entre estudiantes de minorías y de la mayoría, donde el progreso se ha realizado, pero no ha sido en una tasa suficiente para cambiar el rendimiento general de los grupos en un plazo de tiempo razonable (Baker, Griffin y Choi, 2008). A este respecto, en la búsqueda de la equidad en países con población inmigrante, Estados Unidos no está solo en este afán (OCDE, 2005).

Mientras que *No Child Left Behind* (NCLB) ha sido desarrollada con un plan y una velocidad que, de forma comprensible, ha animado a los Estados a seleccionar exámenes que pueden no ser la mejor medida de la eficacia de las escuelas y del aprendizaje de los estudiantes. Lo que necesita un sistema de rendición de cuentas es tanto credibilidad y validez de las medidas empleadas para calificar el progreso hacia los objetivos, como estándares válidos y alcanzables, y, por supuesto, un plan racional para usar los recursos disponibles, financieros y humanos, para alcanzar los resultados deseados.

La validez es una característica muy conocida vinculada a los tests, coincidiendo en la idea general de su ajuste, cualidad técnica y relevancia de los objetivos. Por supuesto, la validez tiene interpretaciones más complejas, la de mayor alcance es la de Messick (1989) que se centra en la calidad de las decisiones adoptadas a partir de los resultados, más que en las características de la prueba en sí misma. Dado que el uso de los tests se ha extendido y sus propósitos nominales se han multiplicado, del mismo modo lo ha hecho la complejidad y la sofisticación requerida de los estudios de validación. Todavía, dado el calendario de desarrollo, se han realizado muy pocos estudios de validez en los Estados Unidos antes de que las evaluaciones se hayan aplicado como medida de la eficacia, y la mayoría de los estudios de validez no abordan el rango completo de propósitos articulados para los exámenes. En lugar de centrarnos en el conjunto entero de propósitos a los que las pruebas de rendición de cuentas (y las evaluaciones formativas) pueden dirigirse, describiremos un conjunto de propósitos que es retador y difícil de abordar. Quizá uniendo experiencias internacionales estemos en condiciones de hacer progresos en estas áreas de validez.

¿Cuáles son los objetivos esenciales de los tests si deben servir como indicadores de alta calidad para la eficacia escolar? En un artículo, publicado hace 15 años,

se enumeran los criterios que las evaluaciones y los tests debían exhibir si querían ser útiles en la escena evaluativa norteamericana (Linn, Baker y Dunbar, 1991; Baker, O'Neil y Linn, 1993). Los criterios relacionan, en parte, los objetivos o estándares que las evaluaciones van a sacar a la luz a raíz de la información recogida en los resultados, ya sea anualmente o en mayores intervalos temporales. Incluimos en esta lista características de complejidad cognitiva: pedir a los estudiantes que realicen procesamientos significativos del contenido y analizarlo utilizando estrategias que requieran múltiples pasos). Una segunda característica relacionada con la riqueza del contenido es que la cognición está sumergida. Este criterio se extendió desde la significación y precisión del contenido como medida. Ambos criterios tuvieron implicaciones sobre cómo el rendimiento tiene que ser juzgado, por ejemplo, si los criterios de puntuación ilustran los niveles de cognición y dominio del contenido deseado. Un tercer criterio fue la equidad, para asegurar que las evaluaciones ofrecían igualdad de oportunidades para estudiantes con distintos contextos socio-culturales o experiencias dentro de los mismos contextos instructivos para demostrar competencia. Inherente a la equidad es evitar la varianza irrelevante para el constructo y el impacto del conocimiento previo, ya sea de los formatos de las pruebas o los contenidos, que sistemáticamente benefician a grupos específicos de estudiantes. La igualdad también se aplicó al acceso y a la capacidad de los estudiantes con necesidades educativas especiales para demostrar competencia en los exámenes. Una interpretación más amplia, implica la extensión de lo que a los estudiantes se les ofrece como razonables oportunidades para aprender el material medido en los test (ver *Raising Standards for American Education, National Council on Education Standards and Testing*, 1992). Una cuarta área para la validez de criterios fue la extensión de lo que el test, o más adecuadamente el sistema de pruebas, aporta a los estudiantes para demostrar transferencia de aprendizajes y generalización de habilidades, estrategias y conocimientos. Estas características son esenciales para asegurar que los estudiantes que han demostrado el uso de procedimientos o resolución de problemas (Vendlinski, Baker y Niemi, 2008) representan una comprensión seria y una expresión en lugar de un producto de procesos superficiales. Mientras que todas las características de la validez son importantes, el resto de este artículo se centra en un conjunto imprescindible si las evaluaciones se usan como base para realizar inferencias sobre la eficacia educativa. El mayor criterio de validez es la «sensibilidad instructiva», componente clave de cualquier razonamiento comprensivo sobre validez para el uso de test en sistemas de rendición de cuentas en educación. La sensibilidad a la instrucción

(Baker, 2008) puede razonarse desde la evidencia de que las puntuaciones de un test (examen o evaluación) están afectadas diferencialmente por la instrucción de alta calidad, idealmente impartida en partes significativas del año escolar. El descriptor de alta calidad de la instrucción compleja debe diferenciarse del entrenamiento o de la práctica directa de preguntas muy similares a las de las pruebas que afectan a las puntuaciones de estos tests. Este matiz tautológico ha sido ya analizado en la investigación relacionada con aptitudes o interacciones entre rasgo y tratamiento (ver por ejemplo Berliner y Cahen, 1973; Cronbach y Snow, 1977; y Tobias, 1976. Una conclusión que procede de la investigación en esta área fue que las medidas de resultados requisen especificidad si quieren ser sensibles a los tratamientos).

Este artículo considera por qué la sensibilidad instructiva debe estar documentada, cuáles son las fuentes para las recomendaciones en los diseños de evaluación para incrementar tal sensibilidad y como resultado deseable, unificar las características de las pruebas y de la instrucción. La sensibilidad instructiva aparece cuando se mantienen aspectos de las siguientes situaciones: 1) hay una gran variación en la implementación de un currículo explícito o de un plan instructivo, 2) hay diferencias sustanciales en la preparación y en la calidad resultante de los profesores, 3) se da a los profesores flexibilidad para adaptar el currículo; y 4) no hay un currículo explícito o plan de estudios más allá de la enumeración de los estándares que deben alcanzarse (cierto en muchos de los entornos de Estados Unidos). De la sensibilidad instructiva se presentará, primero, su importancia, después, una breve revisión de las estrategias pasadas para estudiar la sensibilidad instructiva en las escuelas y, por último, una breve consideración de recomendaciones para futuros estudios de validación.

## **¿Por qué es importante la sensibilidad instructiva?**

Cuando las medidas del rendimiento sirven como criterio principal para el juicio sobre la eficacia educativa en los sistemas de rendición de cuentas, se requiere algún nivel de evidencia que muestre una relación defendible entre las puntuaciones de un test y la calidad de la instrucción o de otras experiencias educativas ofrecidas en la escuela. A menos que exista una vinculación causal entre ambas, el uso de tales tests

como medidas dependientes oscila entre lo cuestionable y lo no garantizado. La dirección de la vinculación no es que las pruebas estén diseñadas para adaptarse a la instrucción, sino más bien que los objetivos estimulen actividades instructivas apropiadas y el aprendizaje consecuente sea «muestreado» por el test. Este mapa de objetivos para la instrucción y para la evaluación es lo que se supone debe ser la alineación entre currículo y evaluación (Baker, 2005), pero es difícil de desarrollar completamente por razones que presentaremos posteriormente.

Se esperan las evidencias de validez para una de las finalidades de cada test (American Educational Research Association, American Psychological Association y National Council on Measurement in Education, 1999). Por ejemplo, la lógica subyacente de muchos sistemas de rendición de cuentas implica más que la finalidad de identificar la eficacia diferencial de las experiencias instructivas de los estudiantes. En definitiva, hay propósitos relacionados con los sistemas de mejoras a través del tiempo, tal y como se muestra en los patrones de objetivos de crecimiento especificados por *Adequate Yearly Progress* (AYP) en la legislación NCLB o en cualquier otra aproximación que permita el estudio del logro incremental de niveles preestablecidos de rendimiento. Para que tales ciclos de mejora ocurran, los resultados de las pruebas deben ayudar a la selección de materiales instructivos, el diagnóstico y la enseñanza de las actuales cohortes de estudiantes (si el tiempo de los resultados lo permite) o, más fácilmente, de las futuras cohortes de estudiantes (en tanto que el sistema mejora a través del tiempo). Por tanto, es sensible para determinar cómo parte de los estudios de validez, las subordinadas pero necesarias partes de evidencia que muestran si los profesores pueden esbozar inferencias a partir de los resultados de los estudiantes, aplicarlas a los déficit de rendimiento de los estudiantes y desarrollar o aplicar un conjunto de estrategias instructivas alternativas para los estudiantes con déficit, problemas de concepto o inadecuaciones. Estas conductas son esenciales para la teoría de la acción subyacente a todos los sistemas de rendición de cuentas que intentan medir los efectos intencionales de la enseñanza y de la escolaridad (Baker y Linn, 2004). La lógica es también la clave de los conceptos de evaluación formativa (Black y Wiliam, 2003; Pellegrino, Chudowsky y Glaser, 2001). Por supuesto, hay una vinculación intermedia crítica en todas estas finalidades, a saber, que la enseñanza realmente produzca aprendizaje. Déjenos considerar la sensibilidad instructiva y sus derivaciones desde el diseño de los iniciales sistemas de evaluación más íntimamente ligados a la enseñanza y a la mejora. Entonces, como contraste, revisaremos el diseño y los objetivos de los actuales sistemas basados en estándares.

## La derivación de las medidas de sensibilidad instructiva

En los años sesenta y setenta, estimulados por un énfasis en los sistemas auto-instructivos, se produjo un auge de las pruebas referidas a criterio (CRT) o su desarrollo, los tests referidos a dominio (DRT). Estas aproximaciones a la medida fueron tratadas como novedad y apoyadas por importantes académicos, muchos de los cuales habían trabajado en el desarrollo de programas de instrucción o de otros sistemas instructivos integrados antes de entrar en el tema de la medida. Las CRT fueron contrastadas con pruebas referidas a la norma –diseños que intentan una distribución normal de los resultados para permitir ordenar las comparaciones de los examinados (Popham y Husek, 1969)–. La denominación de CRT fue acuñada por Glaser (1963), pero constituida en un trabajo previo (Lumsdaine y Glaser, 1960). El desarrollo de CRT fue diseñada por académicos en matemáticas y áreas científicas vinculadas a un conjunto de teorías (Hively, Patterson y Page, 1968) quienes inventaron y explicaron las propiedades y los beneficios potenciales de las pruebas referidas a dominios.

## Cuestiones en el diseño de Pruebas Referidas a Criterio (CRT)

Entre las principales propiedades de las pruebas referidas a criterio está el punto de partida del diseño: un dominio bien especificado de contenido y habilidades intelectuales (ahora orientación cognitiva). Proveniente desde una perspectiva conductista del aprendizaje, el diseño de esos tests estaba muy especificado, y a lo largo del tiempo cambió de lugar desde la simple enumeración detallada hasta incluir un mayor desarrollo y racionalización de las demandas cognitivas. Actualmente, el nivel de especificidad de esas medidas ha hecho más que necesario crear secuencias instructivas realizables y «copiables» en el tiempo. En lugar de empezar con un constructo general y especificaciones del test que distribuía los ítems en uno o más formatos a través de una amplia gama de temas, la nueva alternativa requiere contenidos bien definidos y habilidades delimitadas incluyendo dominios de contenido, problemas tipo, formatos y reglas de puntuación que delimiten el grupo de tareas o ítems. Así como surgieron las ideas, lo hizo la noción probabilística del conjunto de ítems (por ejemplo los subconjuntos difusos), que caracteriza la precisión del conjunto de tareas o la prueba de ítems dentro de las características explícitas del dominio. Se espera que una muestra de tareas

construida de manera extensa y cuidadosa debería medir profunda y directamente tanto dentro del dominio total como de sus diferentes subpartes.

Estos límites del dominio o reglas pretendían también tener un poder «instruccional» porque podían ser transparentes (lenguaje claro y apoyado en ejemplos) y compartidos directamente con los profesores que esperaban controlar la instrucción. Los profesores que desarrollaron o aplicaron una instrucción en la que se exhibían esas características de los dominios explicados esperaban producir una ganancia previsible en el rendimiento en las CRT, al igual que los sistemas de instrucción que se esperaba que fuesen efectivos después de un ciclo o dos de revisión. Dado que el impulso de las CRT vino del diseño de sistemas de instrucción sistemática, se esperaba que los alumnos con menor rendimiento satisfactorio debían conseguir cambios adicionales, con una instrucción más refinada, aprender y de esa forma demostrar sus habilidades con un rendimiento mejorado en los tests. Esta frecuente visión ha caracterizado la formación durante un buen número de años. Esta idea de definir el universo a partir de que las tareas de los tests pueden ser bastante esbozados (Hively et al., 1968) tuvo la mayor parte del poder en la idea de las Pruebas Referidas a Criterio.

## Los informes CRT

Otra característica más destacada y fácil de implementar de las CRT, se centró en los informes. Los resultados de las CRT deberían informar de manera diferente a los tests comerciales disponibles. Las CRT intentaban dar a conocer algunas ideas de las competencias o dominio respecto al (tal vez arbitrario) cuerpo de conocimientos y habilidades en lugar de apuntar a un particular punto dentro de un amplio constructo de capacidad. De esa manera, los informes de las CPR incluyen, al comienzo, medias en bruto como el porcentaje correcto o el porcentaje de logro de algunos objetivos en gran parte arbitrarios (el 90% logrará 90%). Esos números parecen dar una imagen más definitiva de la adquisición de las competencias proyectadas del alumno cuando se contrastan con valores normativos transformados que reflejan el lugar del sujeto en la distribución de los examinados, pero eran fácilmente objeto de manipulaciones. La fijación de puntuaciones de corte con el fin de dividir el grupo en varios niveles de competencia, por si mismo generó una industria artesanal sobre cómo desarrollar puntuaciones de corte y descripciones verbales adicionales que caracterizaran los

resultados de los alumnos. Es de suponer que esos niveles tuvieron que ser validados también. Fue la parte de presentación de informes de CRT la que capturó la atención de desarrolladores y usuarios de los tests. Es relativamente fácil convertir cualquier registro de rendimiento en distribuciones de frecuencia dentro de un particular rango o rangos de puntuaciones, establecidos por un standard de rendimiento o puntuaciones de corte. De hecho, los protocolos de los informes están repletos de requisitos de diseño, que nos sitúan para las prácticas de evaluación a gran escala.

## La situación actual. Evaluaciones con base normativa

Los sistemas de pruebas con base normativa a menudo son entendidos y representados como si fuesen CRT, y es seguramente la impresión dada por la noción de base normativa, por lo menos como se practica en gran parte de los EE.UU. Estos sistemas, de hecho, usan estándares como el origen de su diseño, cuyas descripciones verbales circunscriben únicamente en términos generales las tareas dentro de los ítems. Antes que utilizar las reglas de diseño para el conjunto de ítems, los tests de base normativa han adoptado las características de información de las CRT. Evidentemente, usan criterios de corte entre categorías, como por debajo de básico (*Below Basic*), etc. Estos son fijados por complejos planteamientos relativos a los ítems cuyo propósito podría o no perdurar en procedimientos de escala posteriores. En el diseño frontal, los detalles de la evaluación son menos implícitos. Rara vez apelan a un constructo real validado, consideración que podría ser apropiado a la estructura de su diseño.

La instrucción alineada a estándares y evaluaciones bajo esas condiciones presenta un problema desafiante. Los estándares se piensa que deberían guiar la instrucción y debería ajustarse al centro de la clase. Los tests deberían ser meros indicadores de la adquisición y aplicación de habilidades y conocimientos.

Se pueden encontrar un conjunto de procedimientos (Herman & Webb, 2007) para decidir cuánto ajuste se ha logrado (Baker, 2005). En su mayor parte, estos enfoques se dirigen a la referencia de contenidos que sirven de modelo teórico a ítems relevantes en el test. Aunque se han hecho algunos esfuerzos por establecer el alcance de la medición (Webb, 1999), no es en absoluto un criterio universal y, por tanto, muchas pruebas estatales pueden tener un número relativamente reducido de ítems, o tan pocos que uno o una fracción de uno, localiza un estándar o uno de sus elementos.

La razón para decidir usar estas fuentes de medida incluye el tiempo asignado para la realización de pruebas, la velocidad de ejecución y los costes.

Sin embargo, cuando los estándares no son medidos de forma adecuada (tal vez debido a obstáculos como su número, el tiempo o los costes), ¿qué debería hacer un profesor frente a las sanciones? Es mucho más prudente para ese profesor evitar sanciones mediante la enseñanza de los temas que aparecen en pruebas con cierta frecuencia. Ellos pueden obtener esta información de las pruebas de la inspección, personalmente o por medio de diseños de inspección sustitutos que determinen qué se incluye en realidad en los tests. El resultado es conocido, en la práctica el contenido de la prueba se ha convertido en la norma en muchas escuelas y, en particular, en aquellas con riesgo de sanción. Las consecuencias de este planteamiento son múltiples: los esfuerzos para enseñar estándares se están perdiendo a favor de la enseñanza de contenidos testados; la coherencia y el conocimiento acumulado se pierde; los estudiantes con menor rendimiento tiene poca atención sistemática a subtareas de aprendizaje que podrían sustentar varios estándares y resultados futuros. Las adquisiciones de procedimientos o de trucos en el procedimiento puede ser más fácil de enseñar que conceptos dificultosos en abundantes aplicaciones.

Esto no tiene un gran valor en la esfera política, donde muchos creen que los detalles de test específicos no importan tanto como sus títulos, y los matices, como la profundidad de muestreo, puede ser concebido como una característica técnica pero no esencial. Cuando las medidas están fuertemente correlacionadas, se ha opinado públicamente por los políticos, que son intercambiables unas por otras. Así que, incluso en las evaluaciones con base normativa, muchos de los enfoques tradicionales hacia la tarea o el desarrollo de ítems y el muestreo se utilizan bastante despreocupadamente.

## **Objetivos múltiples y sensibilidad instructiva**

Al igual que ha sucedido con los requisitos para las medidas de rendición de cuentas desarrolladas, ha llegado a ser cada vez más patente que se esperaba que los tests sirviesen para diferentes propósitos, algunos de ellos señalados anteriormente. La mayoría de los expertos en psicometría señalan que cada uno de los objetivos del test es digno de su propio test, con la finalidad de optimizar la calidad de la consiguiente

decisión que debe llevarse a cabo, ya se trate de la entrada, la colocación, o el sistema de monitorización. Parece que las pruebas creadas para un propósito rara vez pueden ser adaptadas para servir a un propósito diferente. Por supuesto, el proceso de adaptación era la única forma en la que las pruebas de rendición de cuentas se podían crear de manera rápida y económica. Sin embargo, si tales pruebas estaban principalmente diseñadas usando especificaciones generales más que dominios bien especificados u ontológicos, ¿cómo puede ser abordado el concepto de sensibilidad de la instrucción?

## ¿Instrucción de alta calidad?

La cuestión es sencilla, ¿cómo saber que las evaluaciones de medida nominales ansían resultados que actualmente son sensibles a una instrucción de alta calidad? Si hemos esbozado las características de las medidas que podrían hacerlas sensibles a la instrucción, ¿cómo debería ser caracterizada la calidad de la instrucción? En primer lugar, y como de costumbre, la instrucción debería asociarse sustantivamente a los objetivos a los que se dirige. Debería representar la clave cognitiva y los elementos de contenido dispuestos en secuencias diseñadas para garantizar condiciones previas. En el mejor de los casos, la instrucción debe proporcionar los principios básicos y las estrategias de apoyo pertinentes y los prerrequisitos de conocimiento, de forma que los estudiantes entiendan por qué están utilizando enfoques particulares y donde encajan en el dominio o en una secuencia de la instrucción. En la medida de lo posible, la instrucción debe guiarse por los resultados de la investigación sobre el aprendizaje que traten de minimizar el peso memorístico o cognitivo (Sweller, 1999), para apoyar el diseño desarrollado (Chi, Glaser, & Farr, 1988), y proporcionar información pertinente. Para apoyar la transferencia, la instrucción debería llevar consigo una amplia gama de formatos para los tests, así como otras estrategias aplicadas e integradas para mostrar los logros.

Éstos incluyen retroalimentación adjunta al dominio práctico adecuado, la comprensión de las intenciones (motivación), la diferenciación individual, y una secuencia gradual de aprendizaje todos implicados en la participación activa de los estudiantes (véase Popham y Baker, 1970, para un tratamiento precoz de estos conceptos en un contexto de preparación del profesor). Estas aplicaciones de los principios pedagógicos son viables y razonablemente eficaces. A pesar de esto, el grado de intensidad,

mezcla, y frecuencia de estos y otros principios de la instrucción, andamiajes, y niveles de engranaje, por ejemplo, podría sugerir refinados estudios de validez. Si podemos ponernos de acuerdo sobre la parte de «calidad» de la instrucción, podemos invertir nuestros análisis y ver qué tipo de prueba responde mejor.

## **Condiciones asumidas para la Sensibilidad de la Instrucción: un resumen provisional**

Para estudiar la sensibilidad de la instrucción, hay algunas cuestiones recurrentes que deben incluirse y deberían responderse afirmativamente, de manera que al menos haya un alto tratamiento de calidad. Estas preguntas son:

- ¿Los resultados de los tests comunican dominios operacionales para la acción instructiva? ¿Son entendidos por los profesores?
- ¿Hay experiencia y tiempo para la enseñanza basada en los resultados?
- ¿Para la atención a las diferencias?
- ¿Pueden los profesores investigar dónde se necesita ayuda (evaluación formativa de alta calidad)?

## **¿Cómo podemos saber? Evidencias a favor y en contra de la Sensibilidad de la Instrucción.**

Las indicaciones anteriores eran una lista de criterios de validez para las evaluaciones (Baker, O'Neil, y Linn, 1993). El criterio de validez especificado en este artículo proporciona gran parte del pegamento intelectual de la agenda de I+D del CRESST para las próximas dos décadas. Las cuestiones claves que podrían ser respondidas por los datos de la inspección estatal, con certeza, nos da una estimación bruta de la probabilidad de encontrar la sensibilidad de la instrucción, son las siguientes:

- ¿Cuál es la relación entre la cantidad de tests prácticos y el tiempo total de instrucción funcional y cómo se conecta con los resultados de las pruebas?
- ¿Hay diferencias en las tasas de exposición a esta práctica por subgrupos?

- ¿Existen tratamientos de instrucción focalizados en la rápida adquisición de prerrequisitos?
- ¿Es evaluado el rendimiento en la transferencia de tareas?
- ¿Existe evidencia en los datos de que el aprendizaje individual es acumulativo y sostenido?

Los niveles actuales de evidencia o bien están vacíos o sólo compete modestamente a la mayoría de nuestros estados.

## Las experiencias del CRESST en cuestiones de Sensibilidad Instructiva

Como parte de nuestra investigación, prometimos dirigir la atención hacia los criterios de validez que más tarde aparecieron en el artículo de Linn et al., y los criterios utilizados como cimiento para vincular los estudios de evaluación del diseño y la validación. CRESST emprendió el desarrollo de una nueva encarnación de CRT en 1990, llamado *Model-Based Assessments* (Baker, 1997, 2007b) anteriormente denominado «Evaluación Cognitiva-Sensitiva». Este fue nuestro intento de desarrollar evaluaciones sensibles a la instrucción, es decir, para reflejar un dominio complejo cuyos rasgos se podrían utilizar para orientar la planificación de la instrucción de los profesores o de otros. Además, hemos tenido preguntas acerca de qué atributos de las evaluaciones podrían interpretarse como un dominio independiente, es decir, que fuese útil a través de los diferentes temas y materias, los cuales se centrasen en criterios extraídos de la literatura sobre el aprendizaje de un determinado dominio *What Students Know* (Pellegrino, Chudowsky, & Glaser, 2001), y si el desarrollo de esas medidas podría centrarse en primer lugar en demandas cognitivas que podrían ser dominios independientes y, por tanto, reducir el costo de elaboración de subsiguientes medidas de alta calidad.

## La medición de la práctica de las aulas

El CRESST, junto con un gran número de investigadores de la comunidad educativa, se ha interesado en la medición de la práctica en las aulas, o en la oportunidad de aprender

(Aguirre-Muñoz *et al*, 2006; Herman & Abedi, 2004; Herman & Klein, 1997; Yoon & Resnick, 1998). Los enfoques utilizados incluyen las observaciones, los autoinformes de conocimientos y pedagógicos (Baker *et al*, 1996), las observaciones (Boscardin *et al*, 2004), las tareas del estudiante y su trabajo (Aschbacher, 1999; Aschbacher & Clare, 2001; Matsumura *et al*, 2002; Matsumura y Pascal, 2003; Matsumura *et al*, 2006), y la interpretación por parte de los maestros de los resultados y los planes de instrucción (Herman & Baker, 2003).

Estos estudios se llevaron a cabo en el CRESST con el fin de determinar en qué medida nuestras evaluaciones aumentaban, o en otras palabras, el grado en que la fidelidad entraba en conflicto con la realidad de la clase (Baker *et al*, 1996). También hemos llevado a cabo estudios de proceso-producto utilizando medidas externas disponibles (Goldschmidt *et al*, 2007) (en el rendimiento de los tests del CRESST y los exámenes de graduación de las escuelas de secundaria), a menudo en el contexto de evaluar una intervención. En estos estudios buscábamos las relaciones entre los autoinformes, la observación, o los artefactos (Matsumura *et al*, 2002; Stecher *et al*, 2007). Todos estos estudios, si bien interesan en cuanto que hay pocas relaciones, casi nunca producen datos convincentes debido a la confusión de la calidad del profesorado y logro del alumno, aspecto señalado por el trabajo de la «razón de ventaja» (*odds-ratio*) de Rogosa (Rogosa, 1999a, b, c).

## Referencias bibliográficas

- AMERICAN EDUCATIONAL RESEARCH ASSOCIATION (2008). *Unleashing the power of formative assessment: A strategy for integrating cognitive research, assessment, and instruction*. Session 13.028 at the annual meeting of the American Educational Research Association, New York.
- AMERICAN EDUCATIONAL RESEARCH ASSOCIATION, AMERICAN PSYCHOLOGICAL ASSOCIATION, & NATIONAL COUNCIL ON MEASUREMENT IN EDUCATION. (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- AGUIRRE-MUÑOZ, Z., BOSCARDIN, C. K., JONES, B., PARK, J. E., CHINEN, M., SHIN, H. S., LEE, J., AMABISCA, A. A. Y & BENNER, A. (2006). *Opportunity to learn measures* (CSE Rep. 678). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing.

- ASCHBACHER, P. R. (1999). *Developing indicators of classroom practice to monitor and support school reform* (CSE Rep. 513). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing.
- BAKER, E. L. (in press). Learning and assessment in an accountability context. In K. Ryan & L.A. Shepard (Eds.), *The future of test-based educational accountability*. Mahwah, NJ: Erlbaum.
- (1997). Model-based performance assessment. *Theory Into Practice*, 36, 247-254.
- (2005). *Aligning curriculum, standards, and assessments: Fulfilling the promise of school reform*. In C.A. DWYER (Ed.), *Measurement and research in the accountability era* (pp. 315-335). Mahwah, NJ: Erlbaum.
- (2007). Model-based assessments to support learning and accountability: The evolution of CRESST's research on multiple-purpose measures. *Educational Assessment* (Special Issue), 12(3&4), 179-194.
- (2008). Empirically determining the instructional sensitivity of an accountability test. Paper presented at the annual meeting of the American Educational Research Association. In W. James Popham, *Empirically Determining the Instructional Sensitivity of an Accountability Test: Alternative Approaches*. New York.
- BAKER, E. L. (2008). Empirically determining the instructional sensitivity of an accountability test. Artículo presentado en la reunión anual de la American Educational Research Association. In W. JAMES POPHAM session 28.072, *Empirically Determining the Instructional Sensitivity of an Accountability Test: Alternative Approaches*. New York.
- BAKER, E., GOLDSCHMIDT, P., MARTÍNEZ, F. & SWIGERT, S. (February, 2002). *In search of school quality and accountability: Moving beyond the California Academic Performance Index (API)* (Deliverable to OERI, Contract No. R305B6002). Los Angeles: University of California, National Center for Research on Evaluation, Standards and Student Testing (CRESST).
- BAKER, E. L., GRIFFIN, N. C. & CHOI, K. (in press). *The achievement gap in California: Context, status, and approaches for improvement*. Paper prepared for the California Department of Education, P-16 Closing the Gap Research Council "Connecting the Dots and Closing the Gap." Davis, CA: University of California, Center for Applied Policy in Education (CAP-Ed).
- BAKER, E. L. & LINN, R. L. (2004). Validity issues for accountability systems. In S. H. Fuhrman & R. F. Elmore (Eds.), *Redesigning accountability systems for education* (pp. 47-72). New York: Teachers College Press.
- BAKER, E. L., NIEMI, D., HERL, H., AGUIRRE-MUÑOZ, Z., STALEY, L., LINN, R. L. & ROGOSA, D. (1996). *Report on the content area performance assessments (CAPA): A collaboration*

- among the Hawaii Department of Education, the Center for Research on Evaluation, Standards, and Student Testing (CRESST) and the teachers and children of Hawaii (Final Deliverable). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing.
- BAKER, E. L., O'NEIL, H. F., J. R. & LINN, R. L. (1993). Policy and validity prospects for performance-based assessment. *American Psychologist*, 48, 1210-1218.
- BAKER, E. L., PHELAN, J., CHOI, K., NIEMI, D., VENDLINSKI, T., GRIFFIN, N., HERMAN, J. L. & HOWARD, K. (in progress). *Design and validation of POWERSOURCE© assessments and instructional materials*. Los Angeles, University of California, National Center for Research on Evaluation, Standards, and Student Testing.
- BERLINER, D. C. & CAHEN, L. S. (1973). Trait-treatment interaction and learning. *Review of Research in Education*, 1, 58-94.
- BLACK, P. & WILLIAM, D. (2003). In praise of educational research: formative assessment. *British Educational Research Journal*, 29, 623-637.
- BOSCARDIN, C. K., AGUIRRE-MUÑOZ, Z., CHINEN, M., LEON, S. & SHIN, H. S. (2004). *Consequences and validity of performance assessment for English learners: Assessing opportunity to learn (OTL) in grad 6 language arts* (CSE Rep. 635). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing.
- CHI, M. T. H., GLASER, R. & FARR, M. (Eds.). (1988). *The nature of expertise*. Hillsdale, NJ: Erlbaum.
- CLARE, L. & ASCHBACHER, P. R. (2001). Exploring the technical quality of using assignments and student work as indicators of classroom practice. *Educational Assessment*, 7, 39-59.
- CRONBACH, L. J. & SNOW, R. E. (1977). *Aptitudes and instructional methods: A handbook for research on interactions*. New York: Irvington.
- GLASER, R. (1963). Instructional technology and the measurement of learning outcomes: Some questions. *American Psychologist*, 18, 519-521.
- GOLDSCHMIDT, P., MARTINEZ, J. F., NIEMI, D. & BAKER, E. L. (2007). Relationships among measures as empirical evidence of validity: Incorporating multiple indicators of achievement and school context. *Educational Assessment* (Special Issue), 12 (3&4), 239-266.
- HERMAN, J. L. & ABEDI, J. (2004). *Issues in assessing English language learners' opportunity to learn mathematics* (CSE Rep. 633). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing.
- HERMAN, J. L. & BAKER, E. L. (2003). *The Los Angeles Annenberg Metropolitan Project: Evaluation findings* (CSE Tech. Rep. No. 591). Los Angeles: University of California,

- National Center for Research on Evaluation, Standards, and Student Testing (CRESST).
- HERMAN, J. L. & KLEIN, D. C. D. (1997). *Assessing opportunity to learn: A California example* (CSE Rep. 453). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing.
- HERMAN, J. L. & WEBB, N. M. (2007). Alignment methodologies. *Applied Measurement in Education*, 20, 1-5.
- HIVELY, W., PATTERSON, H. L. & PAGE, S. H. (1968). A «universe-defined» system of arithmetic achievement tests. *Journal of Educational Measurement*, 5, 275-290.
- LINN, R. L., BAKER, E. L. & DUNBAR, S. B. (1991). Complex, performance-based assessment: Expectations and validation criteria. *Educational Researcher*, 20, 15-21. (ERIC Document Reproduction Service No. EJ 436 999)
- LUMSDAINE, A. A. & GLASER, R. (Eds.). (1960). *Teaching machines and programmed learning: A source book*. Washington, DC: National Education Association of the United States.
- MATSUMURA, L. C., GARNIER, H. E., PASCAL, J. & VALDES, R. (2002). *Measuring instructional quality in accountability systems: Classroom assignments and student achievement* (CSE Rep. 582). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing.
- MATSUMURA, L. C. & PASCAL, J. (2003). *Teachers' assignments and student work: Opening a window on classroom practice* (CSE Rep. 602). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing.
- MATSUMURA, L. C., SLATER, S. C., WOLF, M. K., CROSSON, A., LEVISON, A., PETERSON, M., RESNICK, L. & JUNKER, B. W. (2006). *Using the instructional quality assessment toolkit to investigate the quality of reading comprehension assignments and student work* (CSE Rep. 669). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing.
- MESSICK, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13-103). New York: MacMillan.
- NATIONAL COUNCIL ON EDUCATION STANDARDS AND TESTING. (1992). *Raising standards for American education*. Washington, DC: U.S. Government Printing Office. (ERIC Document Reproduction Service No. ED338721)
- NIEMI, D., WANG, J., STEINBERG, D. H., BAKER, E. L. & WANG, H. (2007). Instructional sensitivity of a complex language arts performance assessment. *Educational Assessment*, 12 (3&4), 215-237.

- NO CHILD LEFT BEHIND ACT OF 2001, Pub. L. No. 107-110, § 115 Stat. 1425 (2002).
- PELLEGRINO, J. P., CHUDOWSKY, N. & GLASER, R. (Eds.). (2001). *Knowing what students know: The science and design of educational assessment*. Washington, DC: National Academy Press.
- POPHAM, W. J. & BAKER, E. L. (1970). *Systematic instruction*. Englewood Cliffs, NJ: Prentice-Hall.
- POPHAM, W. J. & HUSEK, T. R. (1969). Implications of criterion-referenced measurement. *Journal of Educational Measurement*, 6, 1-9.
- ROGOSA, D. (1999a). *Accuracy of individual scores expressed in percentile ranks: Classical test theory calculations* (CSE Tech. Rep. No. 509). Los Angeles: University of California, Center for Research on Evaluation, Standards, and Student Testing.
- (1999b). *Accuracy of Year-1, Year-2 comparisons using individual percentile rank scores: Classical test theory calculations* (CSE Tech. Rep. No. 510). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing.
- (1999c). *How accurate are the STAR national percentile rank scores for individual students? An interpretive guide*. Palo Alto, CA: Stanford University.
- STECHEER, B., BORKO, H., KUFFNER, K. L., MARTINEZ, F., ARNOLD, S. C., BARNES, D., CREIGHTON, L. & GILBERT, M. L. (2007). *Using artifacts to describe instruction: Lessons learned from studying reform-oriented instruction in middle school mathematics and science* (CSE Rep. 705). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing.
- SWELLER, J. (1999). *Instructional design in technical areas*. Camberwell, Australia: ACER Press.
- TOBIAS, S. (1976). Achievement treatment interactions. *Review of Educational Research*, 46, 61-74.
- VENDLINSKI, T. P., BAKER, E. L. & NIEMI, D. (2008). Templates and objects in authoring problem-solving assessments. In E. BAKER, J. DICKIESON, W. WULFECK, & H. F. O'NEIL (Eds.), *Assessment of problem solving using simulations* (pp. 309-333). New York: Erlbaum.
- VENDLINSKI, T. & STEVENS, R. (2000). The use of artificial neural nets (ANN) to help evaluate student problem solving strategies. In B. FISHMAN & S. O'CONNOR-DIVELBISS (Eds.), *Proceedings of the fourth international conference of the learning sciences* (pp. 108-114.). Mahwah, NJ: Erlbaum.
- WEBB, N. L. (1999). *Research Monograph No. 18: Alignment of science and mathematics standards and assessments in four states*. Madison, WI: National Institute for Science Education.

YOON, B. & RESNICK, L. (1998). *Instructional validity, opportunity to learn and equity: New standards examinations for the California mathematics renaissance* (CSE Rep. 484). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing.

## Fuentes electrónicas

BAKER, E. L. (2007a, August/September). The end(s) of testing (2007 AERA Presidential Address). *Educational Researcher*, 36 (6), 309-317. Retrieved October 2, 2007, de [http://www.aera.net/uploadedFiles/Publications/Journals/Educational\\_Researcher/3606/09edr07\\_309-317.pdf](http://www.aera.net/uploadedFiles/Publications/Journals/Educational_Researcher/3606/09edr07_309-317.pdf)

OECD. (2005). *PISA 2003 data analysis manual: SAS® users*. Paris: Author. Available online at <http://www.pisa.oecd.org/dataoecd/53/22/35014883.pdf>

**Dirección de contacto:** Eva L. Baker. Universidad de California en Los Ángeles. National Center for Research on Evaluation, Standards, and Student Testing (CRESST). Los Ángeles California, Estados Unidos. E-mail: [eva@ucla.edu](mailto:eva@ucla.edu)



# Forma básica del crecimiento en los modelos de valor añadido: vías para la supresión del efecto de regresión

## Growth basic form at value-added models: supressing regression effect

María Castro Morera, Covadonga Ruiz de Miguel y Esther López Martín

*Universidad Complutense de Madrid. Departamento de Métodos de Investigación y Diagnóstico en Educación (MIDE).  
Madrid, España*

### Resumen

El presente trabajo estudia los factores que influyen en la pauta de crecimiento de los centros educativos y que están vinculados a las características del diseño longitudinal. En primer lugar, se investiga la inclusión, como predictor de la tasa de crecimiento individual, del nivel inicial de conocimientos de los alumnos. El nivel inicial de conocimientos de los alumnos posibilita el control de los efectos no deseados de la regresión estadística en la clasificación de los centros educativos. Esto es posible gracias a las cuatro ocasiones de medida que incluye el diseño de esta investigación. Ambos problemas se trabajan en un conjunto de resultados de rendimiento en matemáticas de tres cohortes paralelas de un total de 153 centros educativos y 6.689 estudiantes que se corresponden con tres segmentos del sistema educativo español (último ciclo de Primaria y la Educación Secundaria Obligatoria) en el contexto de un diseño longitudinal con cuatro medidas en dos cursos académicos. Los resultados muestran que la no consideración del efecto de regresión provoca una clasificación inadecuada de los centros de enseñanza en los niveles iniciales y en las tasas de crecimiento, si bien es cierto que tiene poco impacto (aunque significativo) en los valores medios que caracterizan al conjunto de los centros. Esta inadecuación en la clasificación de los centros tiene especial relevancia si los índices de valor añadido y tasas de crecimiento se emplean para el establecimiento de planes de mejora o para el reconocimiento, positivo o negativo, de las actuaciones de los centros.

**Palabras clave:** modelos de valor añadido, crecimiento escolar, diseños longitudinales, regresión estadística, función de crecimiento, evaluación de sistemas educativos, sistemas de rendición de cuentas.

### **Abstract**

This paper studies main elements that affects to school growth models linked to longitudinal design features. First, we include student knowledge initial status as predictor of student growth rate. In doing this, it is possible to control unpleasant effects of statistical regression on school rankings. That problem is studied with longitudinal mathematics and reading database of three parallel cohorts that collect data from 153 schools and 6,689 students of three cycles of Spanish educational system (last two grades of Elementary Education and four grades of Compulsory Secondary Education) with four measurement occasions collected at two academic years. Main results show that ignoring regression effect conduct to school misclassifications at school initial status and school growth rate, even when regression effect has small influence at school sample means since its magnitude is low but significant. This misclassification has important repercussions whether these indexes are used to develop improvement programs or to give sanctions or rewards to schools.

**Key Words:** Value-added models, school growth, longitudinal research design, statistical regression, growth function, assessment of educational system, accountability system.

## **Introducción: características de los diseños para el estudio del crecimiento de los centros educativos**

Los diseños de valor añadido ofrecen información valiosa sobre dos cuestiones fundamentales: los niveles de logro en los aprendizajes escolares y las tasas de crecimiento de estos aprendizajes en cada ocasión de medición. Parece claro en la literatura (Zvoch y Stevens, 2003, 2006; Stevens y Zvoch, 2006) que los factores relacionados con las características sociales individuales de los alumnos están más relacionados con los índices de nivel de conocimientos y menos con los índices de crecimiento de los mismos. Sin embargo, los niveles iniciales de conocimiento de los alumnos están estrechamente relacionados con el ajuste de las tasas de crecimiento de los centros.

El nivel inicial de conocimientos y la tasa de crecimiento son los dos elementos fundamentales de los modelos de valor añadido: punto de corte y pendiente. Las

relaciones entre el punto de corte y la pendiente pueden ser variadas. Rogosa (1995) señalaba en su ya clásico artículo centrado en los mitos sobre la investigación longitudinal que caben relaciones negativas, nulas y positivas entre estos dos parámetros. La interpretación de esta relación en cada caso es un indicador de distintas situaciones.

Una relación negativa entre el nivel de partida de los estudiantes y su tasa de crecimiento puede ser un indicador del efecto de regresión hacia la media, un fenómeno que es ampliamente conocido y documentado (Wilder, 1957; Lacey y Lacey, 1962; Nesselroade, Stigler y Baltes, 1980; Jamieson 1995, 1999; Raykov, 1995). La *Ley de los Valores Iniciales* supone de hecho que cuanto más alto es el nivel de partida menor es la tasa de crecimiento posible. Rogosa señala que esta correlación negativa puede ser efecto del diseño, puesto que el número de ocasiones de medición llega a modificar esta relación. Por el contrario, una relación positiva entre niveles de partida y pendientes es conocida como «efecto abanico» (*fanspread effect*) al incrementar, con el paso del tiempo, la varianza de las variables de respuesta. Esta relación positiva es un indicador de llamado «Efecto Mateo» (Merton, 1963; Tu, Gunnell y Gilthorpe, 2008) de acuerdo con el cual, cuanto más alto es el nivel inicial, mayor puede llegar a ser la tasa de crecimiento.

Desde nuestro punto de vista, y a diferencia de lo que sostiene Rogosa, la no identificación de artefactos estadísticos como la regresión estadística puede llegar a tener efectos poco deseables en el marco de los modelos de valor añadido, ya que los centros suelen ordenarse y categorizarse en función de sus medidas de valor añadido y crecimiento medio, por lo que podrían tomarse decisiones administrativas y de mejora sobre la base de una información enmascarada por algunos artefactos estadísticos identificables. Linn (2000) resaltó la importancia de aportar medidas precisas y adecuadas para apoyar la toma de decisiones de alto impacto basadas en la evaluación de los centros.

Parece necesaria la introducción, tanto en los diseños para la medida del cambio, como en los modelos multinivel de crecimiento, de elementos que permitan evitar la influencia de artefactos estadísticos como la regresión hacia la media y hagan posible adquirir un mayor entendimiento de la forma de las funciones de crecimiento. En el siguiente epígrafe, se describe el modelo básico de valor añadido y las posibles mejoras que podrían introducirse para realizar una mejor estimación del crecimiento de las escuelas.

La finalidad de esta investigación es analizar de forma pormenorizada algunos de los factores que afectan a los modelos de valor añadido. De forma específica, nos centraremos en la necesidad de diferenciar el posible fenómeno de la regresión estadística de la estimación del valor añadido de las escuelas y de sus tasas de crecimiento.

## Modelo de crecimiento y sus posibles mejoras

El modelo de valor añadido más sencillo refleja la existencia de una estructura jerárquica de, al menos, tres niveles: el tiempo ( $t$ ), el alumno ( $i$ ) y la escuela ( $j$ ). Las puntuaciones recogidas en distintos momentos de medición se anidan en los alumnos que las producen (mientras aprenden y evolucionan), y que, a su vez, están agrupados en centros escolares. Así,  $y_{tij}$  es el rendimiento en el momento  $t$  de un alumno  $i$  que asiste a la escuela  $j$ . El modelo nulo queda planteado como sigue (ecuación 1).

$$y_{tij} = \beta_{0ij} + \beta_{1ij} (t-t_0) + \varepsilon_{tij} \quad (1)$$

donde  $\varepsilon_{tij} \sim N(0, \sigma_{\varepsilon}^2)$

En este primer nivel, el término  $\beta_{0ij}$  representa el punto de corte en el momento inicial de la medición. Si, como es el caso, la variable de respuesta es el rendimiento,  $\beta_{0ij}$  representa el rendimiento medio esperado del alumno  $i$  en la escuela  $j$  en el momento  $t_0$ , y, por lo tanto, el *nivel inicial de conocimientos* para ese alumno en esa escuela.

Dadas las condiciones de nuestro interés -el valor añadido- este modelo nulo precisa de la inclusión de un término adicional,  $\beta_{1ij}$  vinculado al predictor tiempo. Así,  $\beta_{1ij}$  representa la *tasa de crecimiento esperado en los aprendizajes* del alumno  $i$  en la escuela  $j$  debida al paso del tiempo y  $(t-t_0)$  el paso del tiempo entre mediciones. El término  $\varepsilon_{tij}$ , por su parte, indica el error aleatorio en este primer nivel que muestra las variaciones aleatorias en las puntuaciones dentro de cada ocasión de medición, y muestra una distribución normal con media 0 y varianza constante como supuesto base.

Como se puede observar, este modelo nulo incluye un término referido al nivel de partida en los aprendizajes,  $\beta_{0ij}$ , y otro a la tasa de crecimiento de los mismos,  $\beta_{1ij}$ . Según este modelo, los alumnos y las escuelas pueden presentar variaciones en sus niveles iniciales, así como en la distinta cantidad de aprendizajes logrados a lo largo del tiempo, lo que implica la existencia de variabilidad en los niveles segundo y tercero. Si bien es cierto que se asume que la relación que hay entre el tiempo y el aprendizaje es lineal (y suponemos que positiva), esto supone afirmar que con el paso del tiempo los aprendizajes *siempre* irán creciendo.

El reconocimiento de la variabilidad entre alumnos y escuelas permite desarrollar los subsiguientes submodelos para cada uno de estos niveles. Así, tanto el nivel de los alumnos, como el nivel de las escuelas tienen su propia representación para el nivel inicial de conocimientos y la tasa de crecimiento.

En el segundo nivel, la ecuación (2) muestra la descomposición del estatus inicial del alumno, de tal manera que  $\beta_{0j}$  representa la media en rendimiento de la escuela (j) en el momento de la primera medición, más un componente aleatorio  $\mu_{0ij}$  que muestra el residuo específico de cada alumno con respecto a este valor central de su escuela. En cierto sentido, podríamos llegar a denominar a este componente el «valor añadido del alumno», pues representa la especificidad del nivel inicial de conocimientos de cada estudiante con respecto al conjunto de su escuela.

$$\beta_{0ij} = \beta_{0j} + \mu_{0ij} \quad (2)$$

$$\beta_{1ij} = \beta_{1j} + \mu_{1ij} \quad (3)$$

Del mismo modo, en la ecuación (3), incluye el término  $\beta_{1j}$ , que muestra la tasa de crecimiento medio de la escuela (j), más un término residual,  $\mu_{1ij}$ , que es lo que se aleja de la pendiente de crecimiento de un alumno específico de la tasa de crecimiento general de su escuela.

La parte aleatoria de este submodelo, compuesta por  $\mu_{0ij}$  y  $\mu_{1ij}$ , muestra una distribución normal bivariada, con media 0 y varianzas  $\sigma_{\mu_0}^2$  y  $\sigma_{\mu_1}^2$  respectivamente y covarianza  $\sigma_{\mu_0\mu_1}$ .

Y de forma análoga, en el tercer nivel (escuelas), podemos identificar, en la ecuación (4), cómo el rendimiento inicial medio de una escuela tiene por ahora dos componentes, el nivel medio inicial del conjunto de la población de escuelas ( $\beta_{00}$ ) y la diferencia específica de esa escuela en el momento inicial ( $\mu_{0j}$ ), que es el indicador del *valor añadido de la escuela*. Igualmente, la Ecuación V describe cómo la tasa de crecimiento de una escuela tiene un componente que muestra el nivel de crecimiento medio para el conjunto de escuelas ( $\beta_{10}$ ) y otro que indica la diferencia específica de esa escuela en la tasa de crecimiento ( $\mu_{1j}$ ), que es la tasa de crecimiento específica de ese centro educativo.

$$\beta_{0j} = \beta_{00} + \mu_{0j} \quad (4)$$

$$\beta_{1j} = \beta_{10} + \mu_{1j} \quad (5)$$

La presentación del modelo en tres niveles con submodelos específicos responde a criterios de claridad expositiva, pues el modelo es único, ya que si se substituyen en la ecuación (1) las ecuaciones (2), (3), (4) y (5) queda el modelo general como sigue:

$$y_{tij} = \beta_{00} + \beta_{10} (t-t_0) + \mu_{0ij} + \mu_{0j} + \mu_{1ij} (t-t_0) + \mu_{1j} (t-t_0) + \varepsilon_{ij} \quad (6)$$

Los términos más relevantes conceptualmente del tercer nivel son los residuos del centro vinculados al punto de corte y a la pendiente, pues cuantifican el valor añadido y la tasa de crecimiento de centro, objeto de interés de este artículo. Los estudios sobre crecimiento y valor añadido suelen centrarse en el ajuste en función de los distintos predictores vinculados a las características del alumno, del centro y de la intervención escolar; pues el valor añadido y la tasa de crecimiento se definen como los residuos del modelo.

### **Inclusión del estatus inicial del alumno como predictor para la identificación del efecto de regresión**

El efecto de regresión en los modelos de valor añadido compete a la relación existente entre el nivel de partida y la tasa de crecimiento de los estudiantes. Se puede manifestar a través de la covarianza negativa y de la correlación negativa entre el punto de corte y la pendiente, a través de un descenso de las medias de los centros que regresan a la gran media tras sucesivas ocasiones de medición y a través de una disminución de la varianza de la variable de respuesta entre sucesivas mediciones.

En los estudios de cambio, el efecto de la regresión se hace visible en el nivel del conjunto de la muestra y no dentro de cada centro específico, y está determinado por el nivel de partida de cada alumno. Es, por tanto, el nivel de partida o estatus inicial del alumno el principal indicador del efecto de regresión. Si los alumnos estuvieran aleatoriamente distribuidos en los centros, cada uno de ellos representaría en la práctica una muestra independiente de la población. El efecto de regresión por el que los alumnos de puntuaciones más extremas en la ocasión inicial regresan hacia la media general, tendría un efecto dentro de cada centro, ya que en cada uno de ellos tendríamos la misma probabilidad de encontrar alumnos con puntuaciones en los rangos superiores e inferiores.

Pero, cuando un centro tiene alumnos con puntuaciones extremas, sólo positivas o sólo negativas, en la siguiente medición la media del centro experimentaría también una regresión hacia la media que puede ser confundida con un valor añadido distinto de cero.

No es infrecuente que los alumnos de determinados centros presenten puntuaciones extremas debido a un efecto de selección que está asociado a las características socio-económicas del contexto en el que se encuentra el centro educativo. Por este motivo, es importante considerar cuáles son los dispositivos del modelo que permiten neutralizar el efecto de regresión.

Se propone, por lo tanto, introducir el estatus inicial del alumno como el principal predictor de la tasa de crecimiento. Es definido como la distancia de la puntuación de rendimiento bruto de un estudiante a la media bruta del rendimiento de los alumnos de la muestra en la primera ocasión de medición.

Si este indicador del estatus inicial se introduce en el modelo general de valor añadido como predictor de la tasa de crecimiento del alumno, es posible que se pueda identificar la magnitud del efecto de regresión hacia la media, si fuera significativa. Así quedaría la ecuación (3) que describe la tasa de crecimiento del alumno  $i$  en la escuela  $j$ , al introducir el estatus inicial del alumno como predictor y tras haber sustituido también la ecuación (5):

$$\beta_{1ij} = \beta_{10} + \beta_{11} (\text{EstatusInicial}) + \mu_{1ij} + \mu_{1j} \quad (7)$$

En la ecuación (7) se redefine el significado del punto de corte.  $\beta_{10}$  es ahora la tasa de crecimiento medio para el conjunto de centros cuando el nivel inicial del alumno es igual al de la media de todos los centros en la primera medición. En esta ecuación, el término  $\beta_{11}$  cuantificaría la influencia del estatus inicial del alumno en su tasa de crecimiento. Si integramos esta ecuación en el modelo general planteado en la Ecuación VI, éste quedaría de la siguiente manera:

$$y_{ij} = \beta_{00} + \beta_{10} (t-t_0) + \beta_{11} (\text{EstatusInicial}) (t-t_0) + \mu_{0ij} + \mu_{0j} + \mu_{1ij}(t-t_0) + \mu_{1j}(t-t_0) + \varepsilon_{ij} \quad (8)$$

La ecuación (8) muestra el modelo general de valor añadido que integra el estatus inicial como predictor. El término  $\beta_{11}$  es ahora un término de interacción que cuantifica la relación entre el nivel inicial del alumno y el paso del tiempo. Además, muestra la tasa de crecimiento del alumno cuando se incrementa su nivel inicial un punto con respecto al conjunto de centros. Si este término tuviera un valor negativo sería un indicador del efecto de regresión. Si, por el contrario, tuviera un valor positivo, sería un indicador del denominado efecto Mateo.

La tasa de crecimiento es un factor que cuenta ahora con dos términos ( $\beta_{10} + \beta_{11}$ ) (*EstatusInicial*) y que mostraría la tasa de crecimiento real de un centro teniendo en cuenta el estatus inicial (cuantificándose, como hemos señalado, el efecto de regresión o el efecto Mateo).

Si se comprobara que el efecto de regresión tiene una presencia significativa el modelo planteado en la ecuación (8) debería ser el modelo nulo de partida en el contexto del valor añadido.

La pregunta central de los sistemas de evaluación orientados hacia la mejora está relacionada con la identificación de los ámbitos en los que es posible mejorar. Las escuelas pueden diferenciarse en su nivel de partida y también en su tasa de crecimiento, como muestran las ecuaciones (4) y (5) (Barton y Coley, 1998). Ambas dimensiones pueden considerarse dicotómicas en función de la posición de los centros por encima o por debajo de los valores de estas dos grandes medias ( $\beta_{00}$  y  $\beta_{10}$ ).

Así establecido, los centros se podrían clasificar en función de su posición en función del cruce de estas dos categorías, identificándose cuatro perfiles de centro diferentes:

- Aquellos centros que tienen un estatus inicial por encima de  $\beta_{00}$  junto con una tasa de crecimiento también por encima de la gran media de crecimiento  $\beta_{10}$ . Los mejores centros estarían en esta categoría, pues comienzan alto y crecen alto.
- Por otro lado, estarían los centros que tienen un estatus inicial por debajo de  $\beta_{00}$  junto con una tasa de crecimiento también por debajo de  $\beta_{10}$ . Serían los centros con una situación complicada, su nivel de partida es bajo y no consiguen promover el aprendizaje de sus alumnos.
- Los centros que tienen un nivel inicial positivo (por encima de  $\beta_{00}$ ), pero una tasa de crecimiento negativa (por debajo de  $\beta_{10}$ ).
- Y los que tienen un nivel inicial negativo (por debajo de  $\beta_{00}$ ), pero una tasa de crecimiento positiva (por encima de  $\beta_{10}$ ).

El posible efecto de la regresión hacia la media puede situar inapropiadamente a los centros en estos dos ejes de referencia. Más allá de las implicaciones para los sistemas de evaluación y *accountability*, una mala clasificación de los centros impide la identificación de las líneas de mejora. La pregunta ¿en qué necesitan mejorar los centros: en valor añadido, en la tasa de crecimiento, en ambos? quedaría sin respuesta clara, pues, si no se asumiese un diagnóstico análogo al sugerido en la ecuación (8), se podría estar enmascarando la verdadera situación de los centros debido a un artefacto estadístico ampliamente conocido.

## Diseño de investigación

Los datos utilizados para la realización de este trabajo proceden del Proyecto de I+D con referencia SEC 2003-09742, ya finalizado, y titulado: *El valor añadido en educación y la función de producción educativa: un estudio longitudinal*. Esta investigación

se ha centrado en la evaluación del progreso académico en Matemáticas y Comprensión Lectora para los cursos comprendidos entre 5º de Educación Primaria y 4º de Educación Secundaria Obligatoria a lo largo de cuatro mediciones sucesivas a una muestra aleatoria de centros y estudiantes de la Comunidad de Madrid realizadas en los cursos académicos 2005-06 y 2006-07.

El diseño seguido puede caracterizarse de «multinivel» y longitudinal, previamente descrito.

Se trata de un estudio longitudinal que tomó como línea base la medición realizada en octubre de 2005, siguiendo la serie temporal con las de junio 2006, noviembre 2006 y junio 2007, para observar el crecimiento y la evolución de las cohortes paralelas incluidas en la muestra, con cuatro ocasiones de medida en dos periodos lectivos. De esta forma, se mide a los alumnos al inicio y al final del curso académico durante dos años consecutivos. La unidad de análisis es el estudiante individual, y las unidades del sistema de evaluación serán las escuelas. Se definieron tres segmentos educativos de interés: el último ciclo de Educación Primaria y los dos ciclos de la secundaria obligatoria. Cada uno de estos segmentos educativos constituye una cohorte.

## **Preguntas de investigación**

El objetivo principal de este artículo es la definición y el estudio del crecimiento en los modelos de valor añadido. De forma más específica, se propone:

- Estudiar el impacto del predictor considerado más importante para la explicación de la tasa de crecimiento, es decir, del nivel inicial de partida de cada alumno individual. De esta forma, se pretende analizar el posible enmascaramiento de tasas de crecimiento muy altas o muy bajas con el efecto de regresión hacia la media.

Se analizó la significación de cada uno de los coeficientes y modelos, así como su impacto en la ordenación de los centros atendiendo al valor añadido específico de cada uno y su tasa de crecimiento.

## **Población y muestra**

La población de referencia considerada es el conjunto de alumnos escolarizados en los centros educativos (tanto públicos, como privados y privados concertados) de la Comunidad de Madrid del último ciclo de primaria y en la secundaria obligatoria en

el año académico 2005-06. La población está constituida por un total de 343.746 alumnos agrupados en 1.798 centros de enseñanza en las cinco áreas territoriales (Capital, Norte, Sur, Este y Oeste) en ese mismo año.

La muestra está compuesta por un total de 153 centros de enseñanza distribuidos proporcionalmente en función del tamaño del estrato (titularidad por área territorial y nivel educativo). Se extrajo una muestra estratificada con la titularidad del centro y área territorial como estratos explícitos, y se tomaron los centros como unidades primarias de muestreo. Se determinó un nivel de confianza del 95%, y, para el efecto de diseño, se asumió un valor de 0,16 para la correlación «intraclase», basándose en datos de investigaciones anteriores. La selección se hizo con probabilidad proporcional al tamaño del centro. El total de alumnos resultantes medidos en cuatro ocasiones a lo largo de los cursos académicos 2005-06 y 2006-07 ha sido de 6.689. La Tabla I muestra la distribución de alumnos y centros por cohortes, nivel educativo, titularidad del centro y área territorial.

**TABLA I. Distribución de la muestra**

COHORTE 1		Capital	Norte	Sur	Este	Oeste	TOTAL
Público	Centros	17	6	14	8	5	50
	Alumnos	401	116	385	261	125	1.288
Concertado	Centros	21	1	4	2	2	30
	Alumnos	699	19	143	64	71	996
Privado	Centros	4	2	2	1	3	12
	Alumnos	171	96	45	32	103	447
TOTAL	Centros	42	9	20	11	10	92
	Alumnos	1271	231	573	357	299	2.731

COHORTE 2		Capital	Norte	Sur	Este	Oeste	TOTAL
Público	Centros	8	2	8	4	3	25
	Alumnos	253	75	250	233	108	919
Concertado	Centros	20	1	4	1	2	28
	Alumnos	696	20	120	41	75	952
Privado	Centros	3	2	1	1	1	8
	Alumnos	120	49	20	73	19	281
TOTAL	Centros	31	5	13	6	6	61
	Alumnos	1.069	144	390	347	202	2.152

COHORTE 3		Capital	Norte	Sur	Este	Oeste	TOTAL
Público	Centros	8	2	6	4	4	24
	Alumnos	204	74	171	174	117	740
Concertado	Centros	20	0	4	1	1	26
	Alumnos	579	0	176	17	4	776
Privado	Centros	3	2	2	1	1	9
	Alumnos	83	75	42	65	25	290
TOTAL	Centros	31	4	12	6	6	59
	Alumnos	866	149	389	256	146	1.806

## Variables medidas

Las variables de respuesta utilizadas han sido las puntuaciones obtenidas en las pruebas elaboradas ad hoc y aplicadas a los alumnos de los centros de la muestra. Las puntuaciones se obtuvieron a partir de las estimaciones derivadas de un modelo de medida TRI. Para garantizar que fueran comparables, se utilizó un diseño de pruebas que permitió la equiparación vertical y horizontal de las escalas obtenidas. El diseño posibilitó tanto la medida puntual del rendimiento en cada ocasión, como el estudio del cambio a través del tiempo por medio del procedimiento MML (Máxima Verosimilitud Marginal) implementado en el programa BILOG-MG (Muraki, 1994).

Las puntuaciones son estimadas en escala estandarizada (sobre los datos de la propia muestra) y posteriormente reescaladas a una media de 250 con desviación típica de 50.

El predictor más relevante en este trabajo es el *estatus inicial de conocimientos del alumno*, pues va a permitir controlar el efecto de la regresión hacia la media. Dado que el efecto de la regresión hacia la media se produce en el conjunto de la muestra y no dentro de cada centro escolar, el estatus inicial se ha definido como la distancia entre la puntuación bruta del alumno en la primera medición (octubre de 2005) y la media del conjunto de alumnos en ese mismo momento.

## Modelos de análisis

Para el estudio de las hipótesis sobre la linealidad del crecimiento, así como de la influencia del estatus inicial en las tasas de crecimiento, se precisan modelos que respeten la estructura jerárquica de los datos. En ese sentido y como ya hemos señalado, los modelos jerárquicos lineales son la herramienta metodológica adecuada para resolver estos problemas (Raudenbush, 2001; Raudenbush y Bryk, 2002). Cuando se aplican junto a medidas válidas para diseños con múltiples puntos temporales, estos modelos ofrecen una herramienta adecuada para el estudio de la estructura y la predicción del cambio individual. Una ventaja de esta aportación es que es flexible con respecto a la estructura temporal de las observaciones. Igualmente importante es que los modelos «multinivel» facilitan el estudio de los efectos de las covariables en cada nivel de anidamiento (Laird y Ware, 1982; Bryk y Raudenbush, 1992; Goldstein, 1995; Snijders y Bosker, 1999; Gaviria y Castro, 2005).

Para la estimación y el contraste de los modelos se utiliza el algoritmo EM (*expectation-maximization*), lo que hace imprescindible el empleo de software específico para modelos jerárquicos lineales, en nuestro caso, concretamente, MLWIN (Goldstein, 1993) (<http://www.cmm.bristol.ac.uk/MLwiN>).

Los criterios de adecuación de los modelos seguirán algunas pautas convencionales dentro de la comparación y selección de modelos, incorporándose a éstas algunos criterios de adecuación específicamente desarrollados para este estudio. Así, la significación de parámetros ( $\alpha = 0,05$ ), la comparación de las razones de verosimilitud entre modelos y la aplicación del principio de parsimonia han sido las estrategias globales de evaluación de modelos. Además, se ha estudiado la ordenación de las escuelas producida por los distintos modelos plausibles, utilizando para ello tablas de contingencia y estudios que establezcan una correlación y de caracterización de los tipos de centros.

## Resultados

Se presentan los resultados de Matemáticas en las tres cohortes estudiadas. La segunda cohorte, correspondiente a los cursos de 1º y 2º de ESO, se ha elegido como ilustración para verificar el objetivo principal de este trabajo. Los resultados de la primera cohorte (5º y 6º de Educación Primaria) y de la tercera (3º y 4º de ESO) se presentan a continuación manteniendo la misma estructura.

### Resultados de matemáticas para la segunda cohorte: 1º y 2º de ESO

El modelo nulo de crecimiento lineal (Ecuaciones I a V) es el probado en primer lugar (Tabla II). En cada tabla de resultados, se presenta el valor del coeficiente y entre paréntesis el error típico de estimación del mismo.

La parte fija del modelo muestra que el rendimiento medio para el conjunto de centros en Octubre de 2005 es de 261,9 puntos y que los alumnos crecen 10,04 puntos, en término medio, por cada ocasión de medición. Así, en Junio de 2006, el rendimiento medio de los alumnos será de 271,94 puntos, y se incrementará en 10,04 puntos más en Noviembre de 2006.

**TABLA II. Modelo nulo de crecimiento lineal: Matemáticas, Segunda Cohorte**

M. Lineal		
<b>P. Fija</b>		
Media del nivel de partida inicial (Oct.05)	$\beta_{00}$	261,905 (2,395)
Tasa de crecimiento medio (Lineal)	$\beta_{10}$	10,046 (0,426)
<b>P. Aleatoria</b>		
Varianza entre puntos temporales (Nivel 1)	$\sigma^2_{\epsilon}$	408,252 (8,801)
Varianza entre alumnos en sus niveles de partida (Nivel 2)	$\sigma^2_{\alpha}$	1182,024 (45,802)
Varianza entre alumnos en sus tasas de crecimiento (Nivel 2)	$\sigma^2_{\beta}$	26,534(3,779)
Covarianza entre el nivel de partida y la tasa de crecimiento (Nivel 2)	$\sigma_{\alpha\beta}$	-156,080 (10,952)
Varianza entre centros en sus niveles de partida (Nivel 3)	$\sigma^2_{\gamma}$	295,856 (63,078)
Varianza entre centros en sus tasas de crecimiento lineal (Nivel 3)	$\sigma^2_{\delta}$	7,329 (1,969)
Covarianza entre el nivel de partida y la tasa de crecimiento (Nivel 3)	$\sigma_{\gamma\delta}$	-30,540 (9,413)
<b>Deviance</b>		81.276,700

Sin embargo, lo más significativo de estos resultados es la parte aleatoria del segundo y el tercer nivel. Se observan covarianzas significativas y negativas entre el punto de corte y la pendiente ( $\sigma_{\mu_{0i1}}$  y  $\sigma_{v_{0v1}}$ ). O, lo que es lo mismo, entre el nivel de partida y la tasa de crecimiento (tanto de alumnos como de centros).

El valor negativo de la covarianza señala que altos valores en los niveles de partida de los centros están asociados a bajos valores de las tasas de crecimiento y viceversa. Es decir, que los centros que comienzan en niveles más bajos de matemáticas tienen en general mayores tasas de crecimiento.

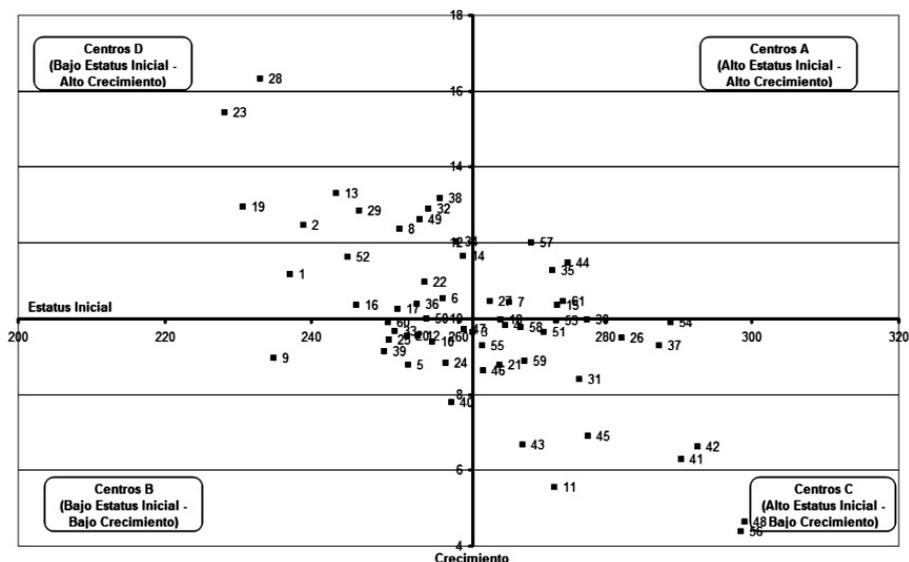
Como es sabido, la correlación entre dos variables (por ejemplo, el nivel de partida y la tasa de crecimiento) es la covarianza estandarizada entre esas dos variables, en el caso de este modelo, obtendríamos una correlación de -0,6558, que es un valor bastante alto.

El valor negativo de la covarianza es, como ya hemos señalado, un indicador del efecto de regresión hacia la media. De hecho, si se observa la distribución de escuelas que muestra la Figura I, se puede ver claramente esta tendencia.

En la Figura I, en el eje de abscisas, está representado el rendimiento medio de cada centro en Octubre de 2005 ( $\beta_{00} + \mu_{0j}$ ) y, en el eje de ordenadas, está representada la tasa de crecimiento de cada centro entre Octubre de 2005 y Junio de 2006 ( $\beta_{10} + \mu_{1j}$ ). Los ejes muestran los valores de  $\beta_{00}$  y  $\beta_{10}$ , que son los valores iniciales, y de tasa de crecimiento medio para todo el conjunto de centros.

Dada la definición y estructura de los modelos de valor añadido, siempre habrá centros situados por encima (denominados «altos») y por debajo (denominados

FIGURA I. Distribución de centros en función de su estatus inicial y tasa de crecimiento según el modelo nulo lineal



«bajos») de la media en cada una de las dos dimensiones. De ahí que se puedan establecer las cuatro categorías de centros en función de su situación inicial y su tasa de crecimiento tal y como anteriormente apuntamos:

- Centros con estatus inicial alto y crecimiento alto (Centros A).
- Centros con estatus inicial bajo y crecimiento bajo (Centros B).
- Centros con estatus inicial alto y crecimiento bajo (Centros C).
- Centros con estatus inicial bajo y crecimiento alto (Centros D).

Desde una perspectiva de intervención optimizada en los centros, cabe preguntarse hasta qué punto el fenómeno que observamos en esta gráfica corresponde a un hecho real o en qué medida es consecuencia de la regresión estadística que aquí se manifiesta.

### Introducción del estatus inicial como predictor en la segunda cohorte

Así, cuando introducimos en el modelo el estatus inicial del alumno como predictor de la tasa de crecimiento del alumno (nivel dos), tal y como se muestra en la Ecuación VIII, obtenemos los resultados que aparecen en la Tabla III.

**TABLA III. Modelo nulo de crecimiento lineal con estatus inicial del alumno: Matemáticas, Segunda Cohorte**

<b>M. Lineal con EI</b>		
<b>P. Fija</b>		
Media del nivel de partida inicial (Oct.05)	$\beta_{00}$	261,884 (2,397)
Tasa de crecimiento medio (Lineal)	$\beta_{10}$	9,845 (0,307)
Tasa de crecimiento en función del nivel de partida: efecto de regresión	$\beta_{11}$	-0,114(0,004)
<b>P. Aleatoria</b>		
Varianza entre puntos temporales (Nivel 1)	$\sigma^2_{\epsilon}$	376,469 (6,656)
Varianza entre alumnos en sus niveles de partida (Nivel 2)	$\sigma^2_{i0}$	1172,411 (39,200)
Varianza entre alumnos en sus tasas de crecimiento (Nivel 2)	$\sigma^2_{i1}$	No significativo
Covarianza entre el nivel de partida y la tasa de crecimiento (Nivel 2)	$\sigma_{i0i1}$	No significativo
Varianza entre centros en sus niveles de partida (Nivel 3)	$\sigma^2_{c0}$	297,453 (62,648)
Varianza entre centros en sus tasas de crecimiento lineal (Nivel 3)	$\sigma^2_{c1}$	3,224 (1,003)
Covarianza entre el nivel de partida y la tasa de crecimiento (Nivel 3)	$\sigma_{c0c1}$	No significativo
<b>Deviance</b>		
		81.257,76

El término  $\beta_{11}$  muestra el impacto que tiene el nivel de partida del alumno con respecto al conjunto de la muestra en la tasa de crecimiento en sus aprendizajes, es decir, representa el ajuste que se debe aplicar a la tasa de crecimiento general debido al efecto de regresión. El peso de este estimador es significativo, pequeño y lógicamente negativo.

Lo más significativo de este modelo es que produce importantes cambios en la parte aleatoria del modelo, reduciendo a cero las covarianzas del nivel dos y tres. Es decir, que la introducción de este término neutraliza la relación existente entre la tasa de crecimiento y el nivel inicial de rendimiento tanto en el nivel del centro, como en el de los alumnos. Es decir, que neutraliza el efecto de regresión detectado. Además, hace también no significativa la pendiente de las tasas de crecimiento en el nivel dos, y conlleva que la pendiente del crecimiento dentro de las escuelas sea la misma para todos los alumnos que asisten a la misma. Según estos resultados, el estatus inicial de los alumnos, tal y como ha sido definido en este estudio, permite identificar y neutralizar los efectos de la regresión estadística hacia la media.

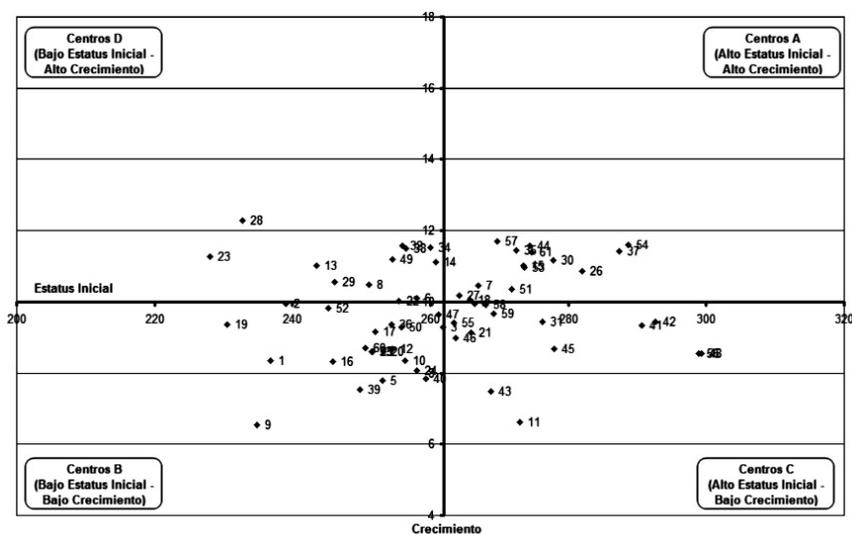
La comparación de las razones de verosimilitud de ambos modelos (Tablas III y IV) muestra diferencias estadísticamente significativas ( $\chi^2 = 18,94$  con dos grados de libertad,  $p = 0,000$ ) a favor de este segundo modelo, que es además más parsimonioso. Por tanto, este modelo, que incluye el estatus inicial como predictor, se perfila como el modelo básico de referencia.

El siguiente aspecto que se debe comprobar es si la ordenación de centros en función del valor añadido y de la tasa de crecimiento es diferente de la ordenación planteada

según el modelo nulo convencional. Si la clasificación fuera diferente, nos interesaría conocer qué centros son los que cambian de categoría.

La Figura II es la análoga a la Figura I una vez identificado el efecto de regresión estadística a través del estatus inicial de alumno. La nube de puntos muestra una distribución que ilustra la no relación entre el nivel de rendimiento en octubre de 2005 y la tasa de crecimiento al final de junio de 2006.

FIGURA II. Distribución de centros en función de su estatus inicial y tasa de crecimiento según el modelo nulo lineal con estatus inicial como predictor



Los centros representados en la Figura I se clasificaron en cuatro categorías correspondientes a cada uno de los cuadrantes definidos por los ejes del plano. Esta clasificación es, por tanto, la que se obtiene cuando se utiliza el modelo definido por las ecuaciones de la I a la V. Al aplicar el modelo de la Ecuación VIII, los mismos centros vuelven a reagruparse en los cuadrantes, tal y como muestra la Figura II.

La tabla de contingencia (Tabla IV) muestra en su diagonal principal la coincidencia de clasificación entre el modelo nulo convencional y el modelo nulo con estatus inicial. Se ve cómo la coincidencia es perfecta en los centros considerados como mixtos en su crecimiento y nivel de partida (centros tipo D y C). Los centros que varían su clasificación son los situados en categorías extremas y homogéneas de acuerdo con su nivel de partida y su tasa de crecimiento (centros tipo A y B).

El modelo con estatus inicial como predictor (denotado como EI en la Tabla IV) sitúa a un total de los seis centros considerados inicialmente como centros tipo D y C en la categoría B que representa a los centros con nivel inicial bajo y crecimiento bajo. Lo que supone una redistribución de aproximadamente el 10% de los centros de la muestra (y del 31,6% sólo en esta categoría).

También es muy llamativo lo que ocurre con los centros tipo C. Nueve centros de esta categoría (el 14,7% de los centros de la muestra) son resituados en la categoría A, mejorando su clasificación.

En conjunto, el modelo de crecimiento lineal que incluye el estatus inicial como predictor redistribuye al 24,5% de los centros de la muestra, lo que, a nuestro modo de ver, es un porcentaje importante dada la trascendencia de algunas de las decisiones que, sobre los centros, se pueden tomar a partir de los resultados de los modelos de valor añadido. Esta reestructuración de la distribución de centros es, además, estadísticamente significativa ( $\chi^2 = 104,052$  con 9 grados de libertad,  $p = 0,000$ ).

TABLA IV. Comparación de distribución de centros: Matemáticas, Segunda Cohorte

			MODELO NULO CONVENCIONAL				TOTAL
			Centros B (BIBC)	Centros C (AIBC)	Centros D (BIAC)	Centros A (AIAC)	
MODELO CON ESTATUS INICIAL	Centros B	Recuento	13	1	5	0	19
	(BIBC)	% con EI	68,4%	5,3%	26,3%	,0%	100,0%
	Centros C	Recuento	0	12	0	0	12
	(AIBC)	% con EI	,0%	100,0%	,0%	,0%	100,0%
	Centros D	Recuento	0	0	14	0	14
	(BIAC)	% con EI	,0%	,0%	100,0%	,0%	100,0%
	Centros A	Recuento	0	9	0	7	16
	(AIAC)	% con EI	,0%	56,3%	,0%	43,8%	100,0%
Total		Recuento	13	22	19	7	61
		% con EI	21,3%	36,1%	31,1%	11,5%	100,0%

A la vista de los resultados, consideramos que el modelo base que debería adoptarse para la introducción de predictores que ayudaran a ajustar la estimación del valor añadido y de la tasa de crecimiento debería ser el modelo que incorpora el estatus inicial del alumno como predictor de la tasa de crecimiento, puesto que ayuda a controlar el efecto de regresión. Dicho efecto tiene una presencia pequeña en el modelo general, por lo que no afecta grandemente a las estimaciones medias del nivel de conocimiento de los centros.

Sin embargo, el efecto de regresión tiene una presencia no deseada desde el punto de vista de la clasificación de los centros educativos, puesto que, gracias al ajuste que

produce el estatus inicial del alumno, se origina una redistribución de algunos centros específicos, los que permite identificar de una manera más ajustada el nivel de partida y la tasa de crecimiento de cada centro.

Por todos los motivos señalados, este modelo es el que debería adoptarse siempre que se den las condiciones en las que sea plausible la presencia del efecto de regresión hacia la media.

A modo de síntesis, podríamos señalar que los resultados apuntan a que el modelo de valor añadido que debe tomarse como referencia para describir el nivel y el cambio en el aprendizaje en el primer ciclo de la educación secundaria es un modelo que incorpora el estatus inicial del alumno como predictor.

### Resultados de matemáticas para la primera cohorte: 5º y 6º de Educación Primaria

Los resultados de la primera cohorte muestran un comportamiento muy similar al de los encontrados en la segunda cohorte. Es decir, se observa la presencia significativa del estatus inicial del alumno como predictor en el segundo nivel, siendo, como hemos venido señalando, un indicador de la presencia del efecto de regresión hacia la media.

La parte fija del modelo lineal para la primera cohorte indica que la media en matemáticas de estos alumnos en Octubre de 2005 era de 251,19 puntos, y que, por su parte, la tasa de crecimiento en cada ocasión de medida fue de 5,7 puntos (Tabla V).

TABLA V. Modelo nulo de crecimiento lineal: Matemáticas, Primera Cohorte

		M. Lineal
<b>P. Fija</b>		
Media del nivel de partida inicial (Oct. 05)	$\beta_{00}$	251,519(1,854)
Tasa de crecimiento medio (Lineal)	$\beta_{10}$	5,700(0,393)
<b>P. Aleatoria</b>		
Varianza entre puntos temporales (Nivel 1)	$\sigma^2_{\epsilon}$	423,803(8,110)
Varianza entre alumnos en sus niveles de partida (Nivel 2)	$\sigma^2_{\alpha_0}$	1354,632(45,801)
Varianza entre alumnos en sus tasas de crecimiento (Nivel 2)	$\sigma^2_{\alpha_1}$	20,778(3,327)
Covarianza entre el nivel de partida y la tasa de crecimiento (Nivel 2)	$\sigma_{\alpha_0\alpha_1}$	-195,406(10,551)
Varianza entre centros en sus niveles de partida (Nivel 3)	$\sigma^2_{\alpha_0}$	249,546(46,318)
Varianza entre centros en sus tasas de crecimiento lineal (Nivel 3)	$\sigma^2_{\alpha_1}$	10,007(2,067)
Covarianza entre el nivel de partida y la tasa de crecimiento (Nivel 3)	$\sigma_{\alpha_0\alpha_1}$	-32,891(8,374)
<b>Deviance</b>		<b>103125,000</b>

Al prestar atención a la parte aleatoria del modelo del segundo y el tercer nivel en los resultados en matemáticas de la primera cohorte, se observan covarianzas significativas y negativas entre el punto de corte y la pendiente ( $\sigma_{\mu_{0i1}}$  y  $\sigma_{v_{0vi}}$ ), como también ocurría en la segunda cohorte.

### Introducción del estatus inicial como predictor en la primera cohorte

Al introducir el estatus inicial del alumno como predictor de la tasa de crecimiento se puede observar cómo el impacto del nivel de partida del alumno con respecto al conjunto de la muestra es de -0,129 puntos (Tabla VI) para esta primera cohorte de alumnos. Respecto a la parte aleatoria de los modelos, podemos observar cómo, al introducir la tasa de crecimiento como predictor, las covarianzas se ven reducidas a cero.

TABLA VI. Modelo nulo de crecimiento lineal con estatus inicial del alumno: Matemáticas, Primera Cohorte

		M. Lineal con EI
<b>P. Fija</b>		
Media del nivel de partida inicial (Oct. 05)	$\beta_{00}$	251,507(1,859)
Tasa de crecimiento medio (Lineal)	$\beta_{10}$	5,568(0,273)
Tasa de crecimiento en función del nivel de partida: efecto de regresión	$\beta_{11}$	-0,128(0,004)
<b>P. Aleatoria</b>		
Varianza entre puntos temporales (Nivel 1)	$\sigma^2_{\epsilon}$	362,927(5,702)
Varianza entre alumnos en sus niveles de partida (Nivel 2)	$\sigma^2_{\mu_{0i}}$	1316,121(38,741)
Varianza entre alumnos en sus tasas de crecimiento (Nivel 2)	$\sigma^2_{v_{1i}}$	0,000(0,000)
Covarianza entre el nivel de partida y la tasa de crecimiento (Nivel 2)	$\sigma^2_{\mu_{0i1}}$	0,000(0,000)
Varianza entre centros en sus niveles de partida (Nivel 3)	$\sigma^2_{\mu_{0i}}$	254,400(46,650)
Varianza entre centros en sus tasas de crecimiento lineal (Nivel 3)	$\sigma^2_{v_{1i}}$	4,079(0,994)
Covarianza entre el nivel de partida y la tasa de crecimiento (Nivel 3)	$\sigma_{\mu_{0i1}}$	-3,176(4,878)
<b>Deviance</b>		103.125,400

Al comparar las razones de verosimilitud de ambos modelos nulos, se observa como se aprecian diferencias estadísticamente significativas en los modelos ( $\chi^2 = 130,9$  con 1 grado de libertad,  $p = 0,000$ ), a favor de este segundo modelo.

Como en el caso anterior, resulta adecuado introducir este predictor para controlar el efecto de regresión. La Tabla VII muestra la reubicación del 29,3% de los centros de la muestra.

**TABLA VII.** Comparación de distribución de centros: Matemáticas, Primera Cohorte

			MODELO NULO CONVENCIONAL				TOTAL
			Centros B (BIBC)	Centros C (AIBC)	Centros D (BIAC)	Centros A (AIAC)	
MODELO CON ESTATUS INICIAL	Centros B	Recuento	11	0	15	0	26
	(BIBC)	% con EI	42,3%	,0%	57,7%	,0%	100,0%
	Centros C	Recuento	0	24	0	0	24
	(AIBC)	% con EI	,0%	100,0%	,0%	,0%	100,0%
	Centros D	Recuento	0	0	18	0	18
	(BIAC)	% con EI	,0%	,0%	100,0%	,0%	100,0%
	Centros A	Recuento	0	12	0	12	24
	(AIAC)	% con EI	,0%	50,0%	,0%	50,0%	100,0%
Total		Recuento	11	36	33	12	92
		% con EI	12,0%	39,1%	35,9%	13,0%	100,0%

### Resultados de matemáticas para la tercera cohorte: 3º y 4º de ESO

A diferencia de las dos cohortes anteriores, en la tercera (para 3º y 4º de ESO) no se observa una covarianza negativa entre el nivel de partida del alumno y su tasa de crecimiento. Si bien es cierto que esto es un indicador claro de la ausencia del efecto de regresión y de la presencia del efecto Mateo. Nos parece interesante mostrar el comportamiento de los modelos en esta tercera cohorte, que ya ha manifestado un comportamiento diferencial con respecto a las dos anteriores, como se puede comprobar en el artículo de Lizasoain y Joaristi (2009) en este mismo monográfico.

En la tercera cohorte, al contrario de lo que ha ocurrido con las otras dos, no hay covarianza estadísticamente significativa en el segundo nivel entre el nivel de partida y la tasa de crecimiento de los alumnos (Tabla VIII). En el tercer nivel, el valor de la covarianza sí resulta significativo y negativo, lo que implica que los centros que comienzan con niveles más bajos en matemáticas tienen mayores tasas de crecimiento y viceversa. En este nivel la correlación entre el nivel de partida de los centros y su tasa de crecimiento es de -0,73, lo que indica un grado de asociación de las dos variables muy elevado.

### Introducción del estatus inicial como predictor en la tercera Cohorte

La introducción del estatus inicial del alumno como predictor de la tasa de crecimiento muestra que el impacto del nivel de partida del alumno con respecto al

**TABLA VIII. Modelo nulo de crecimiento lineal: Matemáticas, Tercera Cohorte**

		M. Lineal
<b>P. Fija</b>		
Media del nivel de partida inicial (Oct. 05)	$\beta_{00}$	266,460(2,764)
Tasa de crecimiento medio (Lineal)	$\beta_{10}$	28,315(1,004)
<b>P. Aleatoria</b>		
Varianza entre puntos temporales (Nivel 1)	$\sigma^2_{\epsilon}$	928,998(19,414)
Varianza entre alumnos en sus niveles de partida (Nivel 2)	$\sigma^2_{\alpha}$	290,125(21,420)
Varianza entre alumnos en sus tasas de crecimiento (Nivel 2)	$\sigma^2_{\delta}$	29,845(5,065)
Covarianza entre el nivel de partida y la tasa de crecimiento (Nivel 2)	$\sigma_{\alpha\delta}$	no significativo
Varianza entre centros en sus niveles de partida (Nivel 3)	$\sigma^2_{\omega}$	402,414(82,482)
Varianza entre centros en sus tasas de crecimiento lineal (Nivel 3)	$\sigma^2_{\nu}$	49,147(10,783)
Covarianza entre el nivel de partida y la tasa de crecimiento (Nivel 3)	$\sigma_{\omega\nu}$	-103,060(25,945)
<b>Deviance</b>		72.038,940

conjunto de la muestra es 0,168 (Tabla IX). Respecto a la parte aleatoria del modelo, al introducir la tasa de crecimiento como predictor, se hace no significativa  $\sigma_{\alpha\delta}$ , es decir, que no hay varianza entre las pendientes de crecimiento de los alumnos que asisten al mismo centro, lo que significa que todos los alumnos del mismo centro crecen de la misma forma.

Sin embargo, se mantiene e incrementa ligeramente el valor de la covarianza  $\sigma_{\omega\nu}$  del tercer nivel, que es un indicador de la relación negativa entre el nivel de partida de los centros y su tasa de crecimiento. Así, los centros con mayor nivel de rendimiento inicial tendrán menor tasa de crecimiento. Esta covarianza no se ha hecho no significativa como en el caso de las otras dos cohortes. Este dato no es, por tanto, un efecto de la regresión estadística, sino que muestra que la posibilidad de que los centros con alto nivel de partida tengan altas tasas de crecimiento es menor. Las explicaciones para este dato son variadas. Por un lado, es posible que, simplemente, los centros con mayor nivel de partida puedan crecer menos en esta etapa en la que aumenta la complejidad curricular de las materias y en la que se incrementan las tasas de fracaso escolar. Por otro, también podría ser un indicador de la falta de sensibilidad de las pruebas diseñadas para medir rendimientos por encima de un determinado punto (el famoso efecto techo). En cualquier caso, parece evidente que son necesarios más estudios que nos ayuden a determinar las causas del comportamiento tan particular de esta cohorte, que se ha manifestado tanto en la estructura de la dimensionalidad de las pruebas (Lizasoain y Joaristi, 2008), como en la de la matriz de varianza-covarianza del rendimiento (Gaviria, Biencinto y Navarro, 2008), ambos en este mismo volumen.

**TABLA IX. Modelo nulo de crecimiento lineal con estatus inicial del alumno: matemáticas, tercera cohorte**

		M. Lineal con EI
<b>P. Fija</b>		
Media del nivel de partida inicial (Oct.05)	$\beta_{00}$	266,410(2,772)
Tasa de crecimiento medio (Lineal)	$\beta_{10}$	28,553(1,080)
Tasa de crecimiento en función del nivel de partida: efecto de Mateo	$\beta_{11}$	0,168(0,006)
<b>P. Aleatoria</b>		
Varianza entre puntos temporales (Nivel 1)	$\sigma^2_{\epsilon}$	946,231(18,278)
Varianza entre alumnos en sus niveles de partida (Nivel 2)	$\sigma^2_{\alpha_0}$	189,774(15,129)
Varianza entre alumnos en sus tasas de crecimiento (Nivel 2)	$\sigma^2_{\alpha_1}$	no significativo
Covarianza entre el nivel de partida y la tasa de crecimiento (Nivel 2)	$\sigma_{\alpha_0\alpha_1}$	no significativo
Varianza entre centros en sus niveles de partida (Nivel 3)	$\sigma^2_{\gamma_0}$	409,266(83,101)
Varianza entre centros en sus tasas de crecimiento lineal (Nivel 3)	$\sigma^2_{\gamma_1}$	59,491(12,571)
Covarianza entre el nivel de partida y la tasa de crecimiento (Nivel 3)	$\sigma_{\gamma_0\gamma_1}$	-129,349(29,495)
<b>Deviance</b>		71.325.420

Al comparar las razones de verosimilitud se observa cómo el modelo que incluye el predictor de estatus inicial de los alumnos es el más adecuado para los datos de la tercera cohorte ( $\chi^2 = 713,52$  con 1 grado de libertad,  $p = 0,000$ ).

Puesto que esta cohorte muestra el Efecto Mateo, no se produce una ordenación distinta de los centros en función del modelo nulo lineal (Tabla VIII), del que incorpora el estatus inicial (Tabla IX). La Tabla X muestra que la clasificación de los centros de acuerdo con ambos modelos es exactamente coincidente. Este resultado es coherente con las tesis sostenidas en este artículo, puesto que, al no haber influencia del estatus inicial en el segundo nivel, no se produce ninguna clasificación errónea de los centros, ya que el modelo no acusa la influencia del efecto de regresión.

**TABLA X. Comparación de distribución de centros: matemáticas, tercera cohorte**

			MODELO NULO CONVENCIONAL				TOTAL
			Centros B (BIBC)	Centros C (AIBC)	Centros D (BIAC)	Centros A (AIAC)	
MODELO CON ESTATUS INICIAL	Centros B	Recuento	12	0	0	0	12
	(BIBC)	% con EI	100,0%	,0%	,0%	,0%	100,0%
	Centros C	Recuento	0	22	0	0	22
	(AIBC)	% con EI	,0%	100,0%	,0%	,0%	100,0%
	Centros D	Recuento	0	0	21	0	21
	(BIAC)	% con EI	,0%	,0%	100,0%	,0%	100,0%
	Centros A	Recuento	0	0	0	4	4
	(AIAC)	% con EI	,0%	,0%	,0%	100,0%	100,0%
<b>Total</b>		Recuento	12	22	21	4	59
		% con EI	20,3%	37,3%	35,6%	6,8%	100,0%

Si tuviéramos que hacer una síntesis para presurosos, diríamos que el estatus inicial de los alumnos, introducido como predictor en el segundo nivel para el crecimiento, es una variable altamente interesante, pues permite cuantificar tanto el efecto de regresión, como el efecto Mateo en las tres cohortes analizadas. Este resultado tiene especial trascendencia pues el efecto de regresión produce una mala clasificación de los centros educativos en función tanto del valor añadido que aportan a sus alumnos, como de la tasa de crecimiento que promueven.

## **Conclusiones**

Este estudio se ha centrado en el análisis de algunos de los factores que afectan a la descripción y la estimación del crecimiento de los aprendizajes escolares utilizando los resultados de matemáticas de una muestra aleatoria y representativa de la Comunidad de Madrid medida en cuatro ocasiones desde octubre de 2005 hasta junio de 2007.

Los resultados de este trabajo de investigación han ilustrado cómo el efecto de regresión tiene una influencia notable en la clasificación de los centros atendiendo tanto a su nivel inicial de conocimientos, como en su tasa de crecimiento. Una vez controlado este efecto es posible estudiar la forma de la función de crecimiento cuestionando y ajustando la hipótesis de la linealidad gracias al diseño planteado, que incluye cuatro momentos de medición.

Como ya hemos señalado, el efecto de regresión se debe a la selección inicial de los alumnos que se agrupan en centros de forma lógicamente no aleatoria. La presencia del efecto de regresión hacia la media implica la presencia en los modelos de valor añadido de una covarianza negativa en el nivel de alumnos entre el punto de corte y la pendiente. El estatus inicial del alumno introducido como predictor de la tasa de crecimiento en el segundo nivel actúa como un neutralizador del efecto de regresión de las medias iniciales de los centros hacia la gran media. Y, si bien es cierto que el impacto sobre los valores medios de la parte del modelo es pequeño, también lo es que reduce la covarianza a valores cero o no significativos, controlando, por tanto, el efecto de regresión.

Esta corrección no tendría trascendencia si los modelos de valor añadido no se utilizaran como una herramienta bien de mejora escolar, bien de rendición de cuentas.

El mayor impacto del control del efecto de regresión se refleja en la clasificación de los centros en función de valor añadido y su tasa de crecimiento, que, por la propia formulación de los modelos, fuerza a situar a los centros al menos por encima o por debajo de esos valores medios. Se ha comprobado que siempre que esté presente el efecto de regresión se produce una discrepancia entre la clasificación de los centros realizada por el modelo que no incluye el estatus inicial del alumno y la clasificación realizada por el modelo que sí lo incluye.

La clasificación correcta de los centros producida por los modelos que identifican y cuantifican el efecto de regresión, utilizada bien para informar internamente, bien para tomar medidas externamente, es un elemento de capital importancia en la configuración y el uso de los modelos de evaluación basados en el valor añadido. De ahí que desde estas páginas se recomiende la introducción del estatus inicial como predictor de las tasas de crecimiento, siempre que se tenga evidencia de esa relación negativa entre el nivel de partida y la pendiente de la función de crecimiento.

Por último, nos gustaría destacar que los valores medios de rendimiento y las tasas de crecimiento medio informan de diferentes dimensiones de la actuación de los centros escolares. Parece, por tanto, importante que los sistemas de evaluación basados en modelos de valor añadido informen y estudien ambos tipos de medida de forma sistemática.

## Referencias bibliográficas

- BARTON, P. Y COLEY, R. (1998). *Growth in school: Achievement gains for the fourth to the eighth grade*. Princeton, NJ: Educational Testing Service.
- BRYK, A. S. & RAUDENBUSH, S. W. (1992). *Hierarchical linear models: applications and data analysis methods*. Newbury Park, CA.: Sage.
- GAVIRIA, J. L., BIENCINTO, C. Y NAVARRO, E. (2009). Invarianza de la estructura de covarianza de las medias de rendimiento académico en los estudios longitudinales en la transición de Educación Primaria a Secundaria. *Revista de Educación*, 348
- GAVIRIA, J. L. Y CASTRO, M. (2005). *Modelos jerárquicos lineales*. Madrid: La Muralla-Hespérides.
- GOLDSTEIN, H. (1995). *Multilevel Statistical Models*. London: Edward Arnold.
- GOLDSTEIN, H. & SPIEGELHALTER, D. J. (1996). League tables and their limitations: Statistical issues in comparisons of institutional performance. *Journal of the Royal Statistical Society, Series A*, 159, 384-443.

- JAMIESON, J. (1995). Measurement of Change and the Law of Initial Values: A Computer Simulation Study. *Educational and Psychological Measurement*, 55 (1), 38-46.
- (1999). Dealing with baseline differences: two principles and two dilemmas. *International Journal of Psychophysiology*, 31, 155-161.
- LACEY, J. I. & LACEY, B. C. (1962). The law of initial value in the longitudinal study of autonomic constitution: reproductibility of autonomic responses and response patterns over a four year interval. En W.M. Wolf (ed.), *Rhythmic functions in the living system. Annals of the New York Academy of Science*, 98, 1257-1290.
- LAIRD, N. M. & WARE, J. H. (1982). Random-effects models for longitudinal data. *Biometrika*, 65, 581-590.
- LINN, R. (2000). Assessments and accountability. *Educational Researcher*, 29, 4-16.
- LIZASOAIN, L. Y JOARISTI, L. (2009). Análisis de la dimensionalidad en modelos de valor añadido: estudio de pruebas de matemáticas empleando métodos no paramétricos basados en TRI. *Revista de Educación*, 348.
- MERTON, R. K. (1963). The Matthew effect in science. *Science*, 199, 55-63.
- NESSELROADE, J. R., STIGLER, S. M., & BALTES, P. B. (1980). Regression toward the mean and the study of change. *Psychological Bulletin*, 88, 622-637.
- RAUDENBUSH, S. W. (2001). Comparing personal trajectories and drawing causal inferences from longitudinal data. *Annual Review of Psychology*, 52, 501-525.
- RAUDENBUSH, S. W. & BRYK, S. A. (2002). *Hierarchical linear models. Applications and data analysis methods*. London: SAGE.
- RAYKOV, T. (1995). On statistical approaches to the study of the «Law of Initial Values». *Quality and Quantity*, 29, 251-271.
- ROGOSA, D. R. (1995). *Myths about longitudinal research*. In J. M. GOTTMAN (ed.). *The analysis of change* (4-66). Mahwah, N.J.: Lawrence Erlbaum.
- SNIJDERS, T. & BOSKER, R. (1999). *Multilevel Analysis*. Thousand Oaks, CA: Sage.
- (2006). Issues in the implementation of longitudinal growth models for student achievement. In R. W. LISSITZ, *Longitudinal and value added models of students' performance* (170-208). Maple Grove, MN: JAM Press.
- WILDER, J. (1957). The lay of initial value in neurology and psychiatry. *Journal of Nervous and Mental Disease*, 125, 73-86.

## Fuentes electrónicas

TU, Y. K., GUNNELL, D. Y GILTHORPE, M. S. (2008). Simpson's paradox, Lord's paradox and Suppression Effects are the same phenomenon the reversal paradox. *Emerging Themes in Epidemiology*, 5 (2), 1-9. Consultado de:

<http://www.ete-online.com/content/5/1/2>

ZVOCH, K. & STEVENS, J. (2003). A multilevel, longitudinal analysis of middle school math and language achievement. *Educational Policy Analysis Archives*, 11 (20). Consultado de: <http://epaa.asu.edu/epaa/v11n20/>

— (2006). Successive student cohorts and longitudinal growth models: an investigation of elementary school mathematics performance. *Educational Policy Analysis Archives*, 14 (2). Consultado de: <http://epaa.asu.edu/epaa/v14n2/>

**Dirección de contacto:** María Castro Morera. Universidad Complutense de Madrid. Facultad de Educación. Departamento de Métodos de Investigación y Diagnóstico en Educación. Avenida del Rector Royo Villanova s/n. 28040 Madrid. E-mail: [maria.castro@edu.ucm.es](mailto:maria.castro@edu.ucm.es)

# Sensibilidad de las especificaciones del modelo de valor- añadido: midiendo el estatus socioeconómico

## Sensitivity of value-added model specifications: measuring socio-economic status

M<sup>a</sup> Eugenia Ferrão

Universidade da Beira Interior. Departamento de Matemática. Covilhã, Portugal

### Resumen

El artículo describe el grado en que dos medidas diferentes de estatus socioeconómico (ESE) o la exclusión de esta variable de control del modelo de valor añadido (VA) cambia las estimaciones del valor añadido de la escuela.

El modelo estadístico usado para estimar el valor añadido de la escuela es el de componentes de la varianza, donde los alumnos son las unidades de nivel uno y las escuelas las unidades de nivel dos. Se incluye también el rendimiento previo como variable explicativa. Los datos utilizados en el artículo proceden de un proyecto de investigación sobre efectividad de las escuelas (*Eficacia Escolar no Ensino da Matemática, 3EM Project*) y fueron recogidos en Cova da Beira (NUT III), Portugal. Es un conjunto de datos longitudinales que permite seguir a los estudiantes a lo largo de su escolarización. No obstante, para el propósito de este artículo, solamente se utilizaron los datos recogidos al comienzo y al final del año académico 2005-06 de los cursos 1º, 3º, 5º, 7º y 8º. Las variables de ESE consideradas son: elegibilidad del estudiante para beca de comedor y libros; y educación de los padres (del progenitor responsable del estudiante). La evidencia presentada parece sugerir que a lo largo de la trayectoria escolar, el rendimiento previo condensa los efectos del ESE. Estos resultados sugieren que los modelos de VA son sensibles a las formas como se mide el rendimiento de los estudiantes, incluso cuando solamente hay una materia.

*Palabras clave:* modelos de valor añadido, nivel socio-económico, ajuste de covariables, rendición de cuentas, sistemas de evaluación de sistemas educativos.

### **Abstract**

The paper describes the extent to which two different measures of socio-economic status (SES) or the exclusion of that controlling variable from the value-added model (VAM) changes the estimates of school value-added.

The statistical model used to estimate the school value-added is a variance component model where pupils are level one units and schools are level two units. Prior achievement is included as explanatory variable as well. The data used in this paper is derived from a school effectiveness research project (*Eficácia Escolar no Ensino da Matemática, 3EM* project) and was collected in Cova da Beira (NUT III), Portugal. It is a longitudinal data set which allows pupils to be followed through their schooling. However, for the purpose of this paper, we only used the data collected at the beginning and at the end of the academic year 2005-06 for the 1<sup>st</sup>, 3<sup>rd</sup>, 5<sup>th</sup>, 7<sup>th</sup> and 8<sup>th</sup> grades. SES variables considered are: (1) the student eligibility for Free School Meals and Books (FSM); (2) Parent's education (parent responsible for the pupil). Evidence shows that prior achievement shrinks ESE effects. Results also show that value-models are sensible to way achievement is measured, even when there is only one subject.

*Key Words:* Value-added models, socio-economical status, adjusting covariates, accountability, assessment systems.

## **El estatus socioeconómico (ESE) en la literatura de los modelos de VA**

Con el término de modelo de valor-añadido (VA) se hace referencia a una familia de modelos estadísticos que se emplean para hacer inferencias sobre la efectividad de unidades educativas, normalmente escuelas y/o profesores (Braun y Wainer, 2007). Educadores, investigadores y responsables políticos generalmente están de acuerdo en que la escolarización es solo uno de los muchos factores que afectan al rendimiento y aprendizaje de los estudiantes. Otro de los factores que tradicionalmente se ha reconocido que contribuye al progreso educativo del estudiante es su estatus socioeconómico, que es un importante predictor del rendimiento. Los sociólogos usan el término ESE para hacer referencia a la posición relativa de una familia o individuo en una estructura social jerárquica, basada sobre su acceso a, o control sobre, bienestar, prestigio y poder. La posición de ESE en esta jerarquía afecta a sus oportunidades educativas y normalmente se utiliza una medida de ESE como variable de control en los modelos de VA o de efectos de la escuela.

Thomas y Mortimore (1996) compararon cinco modelos «multinivel» de complejidad variable para elegir el mejor modelo de VA y seleccionaron aquellos cuyas variables fueran el rendimiento previo de los estudiantes en tests de capacidades cognitivas verbales y cuantitativas no verbales junto con género, edad, etnia, movilidad y beca de comedor. Encontraron que las medidas de rendimiento previo fueron el factor más importante a controlar y Gray, Jesson, Goldstein, Hedger y Rasbash (1995) coincidieron en resultados similares. En la misma línea, Sammons, Thomas y Mortimore (1995) mostraron que el rendimiento previo es el factor más importante requerido para controlar las diferencias de entrada en las medidas de VA (usando una única muestra de la zona del centro urbano de Londres) y también mostraron que la inclusión de factores socioeconómicos en el análisis es muy relevante.

Rubin, Stuart y Zanutto (2004) apoyan la inclusión del ESE como una variable de control en los modelos de VA:

«Así vemos que existen muchas complicaciones cuando se piensa en un experimento aleatorizado ideal e incluso surgen más cuando se considera un estudio que usa datos observacionales. Con los datos observacionales una meta clave es encontrar unidades de tratamiento y control, que sean lo más similares posible en las covariantes del contexto social. Si los grupos parecen muy diferentes en estas covariantes, los resultados probablemente se basen en supuestos del modelo no verificables y en la extrapolación. [...] Puesto que los valores de “porcentaje de minorías” y “porcentaje en pobreza” difieren mucho en diferentes escuelas, como ilustran [...] Tekwke et al. (2004), es probable que los estimadores que ajustan por estas covariantes usando modelos se basen en gran medida en la extrapolación, incluso si los estudiantes fuesen aleatoriamente asignados a estas escuelas después de ser subclasificadas en bloques (con probabilidades radicalmente distintas de asignación de tratamiento entre bloques pero probabilidades similares intra bloques). Esta situación implica una sensibilidad extrema a los supuestos de estos modelos. Si la escuela A no tiene estudiantes que “parezcan similares” a los de otras escuelas, es imposible estimar el efecto de la escuela A con relación a las escuelas de comparación sin hacer supuestos heroicos» (p.109).

Hay algunas experiencias de modelos de VA que no incluyen tales variables (Ladd y Walsh, 2002) o que concluyen que el ESE y las variables demográficas en el nivel del estudiante tienen poco efecto sobre las evaluaciones del valor añadido de los profesores,

puesto que la historia longitudinal del rendimiento de un estudiante sirve como un sustituto de las variables omitidas (Ballou, Sanders y Wright, 2004).

Ladd y Walsh (2002) no controlan el ESE en las estimaciones del VA. Proponen la inclusión de más de un año del rendimiento previo como variable instrumental para ajustar el error de medida en un modelo de desarrollo. Sin embargo, repetidamente se refieren a la influencia e importancia de este constructo para conseguir estimaciones más fiables de VA de la escuela.

«[...] muchas de las escuelas que dan servicio a estudiantes de bajo rendimiento (que también tienden a ser las que atienden a estudiantes con bajo estatus socioeconómico) habrían sido declaradas más efectivas de lo que parecían en las clasificaciones del estado y lo inverso sería cierto para las escuelas que sirven a estudiantes de alto rendimiento (p.11). [...] El resultado combinado podría ser que los profesores y administradores de alta preparación tratarían de evitar las escuelas que atienden a los estudiantes de bajo ESE en favor de las escuelas que atienden a estudiantes de alto ESE. Aunque la evidencia anecdótica de North Carolina es consistente con esta visión, no conocemos estudios sistemáticos sobre la magnitud de este efecto y creemos que debería ser objeto de más investigación. Cuanto mayores son los incentivos, el sistema de rendición de cuentas reduciría más la calidad de la educación en las escuelas en las que las ganancias de rendimiento son más necesarias» (p.16).

McCaffrey, Lockwood, Koretz y Hamilton (2003), a pesar de estar a favor de la inclusión del ESE como una variable contextual del estudiante, concluyen que controlar el ESE y otros factores demográficos solamente en el nivel del estudiante, no basta para eliminar los efectos de las características contextuales en todos los sistemas escolares, especialmente en aquellos sistemas que atienden a estudiantes heterogéneos.

## Medidas de ESE y el modelo

Para el propósito de este artículo se usaron dos variables como indicadores (*Proxy*) del ESE: la educación de los padres y la elegibilidad del estudiante para beca de comedor y libros (BCL). Estas variables son usadas a menudo en los estudios de VA o de

efectos de la escuela/profesor. No estamos completamente seguros de que estas variables sean *válidas* para representar el constructo ESE. Se está trabajando en el desarrollo de un índice compuesto para el estatus socioeconómico y cultural de los estudiantes que incluye: educación y ocupación de los padres; capital cultural (cuántas veces en el año anterior asistió el estudiante a un concierto, fue a un museo, galería de arte, teatro, etc.); y capital social («obligaciones, expectativas y confianza», Coleman, 1988). La gran cantidad de datos necesarios para obtener el índice de ESE implica que, solamente por la inclusión del índice, la cantidad de datos perdidos aumentaría en un 26%, lo que constituye una importante limitación para su uso.

Consideramos un modelo de componentes de la varianza de dos niveles con los estudiantes (indexados por  $i$ ) en el nivel uno y las clases-escuelas (indexadas por  $f$ ) en el nivel dos. Así, el VA se cuantifica mediante los residuos ajustados ( $\hat{u}_{os}$ ) de la ecuación de nivel dos;  $\hat{u}_{os}$  representan la desviación del rendimiento de la clase/escuela ( $\hat{\beta}_{os}$ ) de la media general ( $\hat{\gamma}_{00}$ ), ajustando el rendimiento previo del estudiante ( $x1_{js}$ ) y el ESE del estudiante y de la escuela ( $x2_{js}$  y  $x3_{js}$ , respectivamente). El modelo para la población que deseamos estimar, se escribe como:

$$\begin{aligned} y1_{js} &= \beta_{0s} + \beta_1 x1_{js} + \beta_2 x2_{js} + \beta_3 x3_{js} + \varepsilon_{js} \\ \beta_{0s} &= \gamma_{00} + u_{0s} \\ \varepsilon_{js} &\sim N(0, \sigma_\varepsilon^2) \\ u_{0s} &\sim N(0, \sigma_{u0}^2) \end{aligned}$$

La variable respuesta es la puntuación normalizada en matemáticas (puntuación dos) vinculada<sup>1</sup> (*linked*) con la puntuación de rendimiento previo en matemáticas (puntuación uno). Los datos utilizados en el artículo se obtuvieron de un proyecto de investigación sobre efectividad de las escuelas (*Eficácia Escolar no Ensino da Matemática, proyecto 3EM*) y fueron recogidos en Cova da Beira (NUT III), Portugal. La población objetivo la constituyen los estudiantes matriculados en enseñanza obligatoria (Primaria -cuatro años-, Elemental -dos años-, y Secundaria Inferior -tres años-). La muestra aleatoria es representativa en el nivel del área y de la región NUT III (Vicente, 2006). La muestra inicial fue sobremuestreada para tener en cuenta el posible no consentimiento de los padres, abandono y mortandad experimental, que son un problema conocido en los estudios longitudinales. La mayor tasa de abandono fue del 4,8% en el 8<sup>o</sup> curso. En las clases de educación primaria la tasa fue inferior al 1%. La

<sup>1)</sup> Vinculación (*linking*) por el procedimiento de ítems comunes.

actuación de los profesores y de los directores contribuyó en gran medida a mantener las tasas en un bajo nivel.

Los abandonos y casos perdidos, que se debieron principalmente a la variable educación de los padres, redujeron el número de casos en un 6,1%, 5,7%, 10%, 8,1% y 10,3%, en cada curso, respectivamente. Para el propósito de la estimación de parámetros, los casos perdidos fueron considerados «perdidos al azar» (Little y Rubin, 2002). El diseño de la encuesta fue longitudinal, lo que permitió seguir a los estudiantes durante sus años de escolarización y consistió en tres oleadas: 2005-06, 2006-07 y 2007-08, y los datos se recogieron al comienzo y al final de cada curso académico. Para el propósito de este artículo, usamos solamente los datos recogidos en el curso académico 2005-06 en los cursos 1<sup>o</sup>, 3<sup>o</sup>, 5<sup>o</sup>, 7<sup>o</sup> y 8<sup>o</sup>.

## Resultados

En la Tabla I se presenta el número de unidades estadísticas utilizadas en el análisis y en la Tabla II algunos estadísticos descriptivos de las variables de ESE, como la proporción de estudiantes elegibles para beca de comedor y libros (BCL), la desviación típica de la proporción para el conjunto de las escuelas (una medida de la heterogeneidad del ESE entre escuelas), la desviación típica de la educación de los padres por escuela (una medida de la heterogeneidad del ESE entre escuelas).

TABLA I. Composición de la muestra 1<sup>a</sup> oleada

Curso	Número de:		
	Alumnos	Clases	Escuelas
1	309	35	35
3	327	37	37
5	306	19	9
7	287	18	11
8	248	16	11

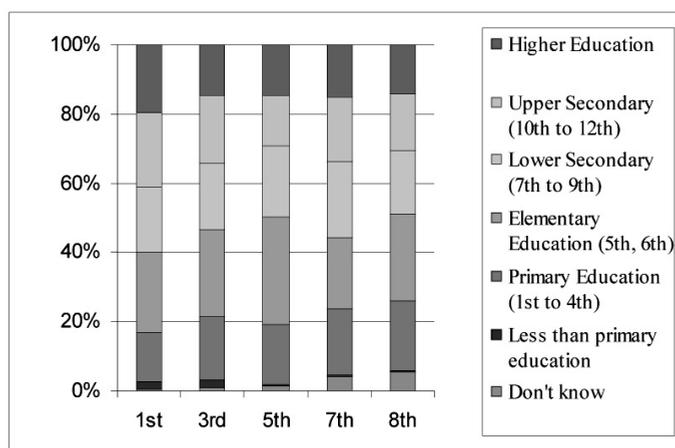
TABLA II. Estadísticos descriptivos de ESE

Curso	Proporción P(BCL=5)	FSM DT (Promedio de la proporción de BCL de la escuela)	Desviación típica de la educación de los padres (promedio de la escuela)
1	0.19	0.16	0.54
3	0.13	0.12	0.57
5	0.39	0.20	0.49
7	0.31	0.13	0.49
8	0.33	0.17	0.46

Comparando la distribución de probabilidad de la BCL en educación primaria con las de educación elemental y secundaria inferior se pueden observar valores muy diferentes, por lo que parece improbable que sean precisos, teniendo en cuenta que es la misma

población de base en términos de la distribución del ESE. Mientras que en la educación primaria la responsabilidad y gestión de los fondos de ayudas sociales a los estudiantes depende del gobierno local (autarquía), en los niveles de educación elemental y superiores, la responsabilidad y gestión depende de cada escuela. Los criterios y los recursos son diferentes en cada subsistema de educación. Así, la BCL parece que es una medida de ESE con error, lo que normalmente se conoce como error de clasificación. Es necesario más trabajo e investigación para ajustar esta clasificación incorrecta. Ferrao y Goldstein (2008) muestran el impacto del error de medida en los estimadores de VA.

FIGURA I. Distribución de la educación de los padres



La Figura I muestra la distribución de la educación de los padres de los estudiantes. El aumento de la educación de los padres entre los cursos primero y quinto es claramente visible en las categorías de secundaria no obligatoria y universidad. En el quinto curso estas categorías representan aproximadamente el 30% mientras que en el primer curso representan el 41%.

## Valor añadido: Estimaciones de parámetros

Las Tablas en el anexo A presentan los estimadores de parámetros del modelo de VA, la especificación del modelo (uno) con diferentes conjuntos de variables de control:

	X2, ESE de los estudiantes <sup>2</sup>	X3, ESE de la escuela
<b>Modelo 0</b>	---	---
<b>Modelo 1</b>	Educación de los padres	---
<b>Modelo 2</b>	Educación de los padres	Educación promedio de los padres
<b>Modelo 3</b>	BCL	---
<b>Modelo 4</b>	BCL	Proporción de BCL

Los parámetros fijados del «modelo 0» son todos estadísticamente significativos ( $p \leq .05$ ) y sus estimadores muestran la fuerte correlación entre la variable respuesta (puntuación dos) y el rendimiento previo (puntuación uno).

La proporción de varianza explicada por el «modelo 0» (puntuación uno) es de 19%, 26%, 48%, 34% y 19%, respectivamente en cada curso. Esto confirma la importancia del rendimiento previo en el modelo de VA. Especialmente en algunos cursos, el rendimiento previo está moderadamente correlacionado con la educación de los padres<sup>3</sup>, por ejemplo en el tercer curso es de -0,33, en quinto es -0,38 y en séptimo es -0,34.

Con respecto al efecto de la educación de los padres sobre las puntuaciones en matemáticas, los resultados del modelo uno muestran una relación negativa ( $p \leq .01$ , para el octavo curso) con la excepción del séptimo curso. El coeficiente de determinación es del 52% para el quinto curso.

El modelo dos, que incluye la variable contextual de educación de los padres no añade otros resultados relevantes, a no ser que el parámetro fijado relacionado a la variable contextual, no es estadísticamente significativo.

Los estimadores de parámetros del modelo tres sugieren que la BCL es solamente estadísticamente significativa en el quinto curso. Anteriormente se ha mencionado la falta de fiabilidad de la BCL como una variable de ESE, desarrollaremos más trabajo sobre esta cuestión antes de comentar los resultados. Lo mismo se aplica al modelo cuatro en el que los resultados sugieren que la variable contextual basada en la BCL solamente es estadísticamente significativa en los cursos primero ( $p \leq .05$ ) y quinto ( $p \leq .01$ ).

<sup>(2)</sup> Educación de los padres- escala invertida y estandarizada

<sup>(3)</sup> Escala invertida

## Comparación de estimadores de VA

Los residuos de nivel dos (estimadores de VA) y los rangos (ordenaciones) producidos por todos los modelos fueron comparados para cada curso. Las Matrices I y II muestran las correlaciones entre los estimadores en el primer curso y la correlación entre los rangos. Entre los modelos que incluyen la variable ESE la correlación es menor (aunque con valores muy altos) en los modelos uno, dos y cuatro.

MATRIZ I. Correlación entre estimadores de VA – Primer curso

	Mod5	Mod4	Mod3	Mod2	Mod1	Mod0
Mod5	1.0000					
Mod4	0.9875	1.0000				
Mod3	0.9201	0.9311	1.0000			
Mod2	0.9028	<b>0.8881</b>	0.9818	1.0000		
Mod1	0.9032	<b>0.8875</b>	0.9795	0.9998	1.0000	
Mod0	0.8821	0.9106	0.9685	0.9843	0.9818	1.000

MATRIZ II. Correlación entre rangos – Primer curso

	Mod5	Mod4	Mod3	Mod2	Mod1
Mod5	1.0000				
Mod4	0.9801	1.0000			
Mod3	0.9157	0.9273	1.0000		
Mod2	0.8940	<b>0.8856</b>	0.9742	1.0000	
Mod1	0.8918	<b>0.8806</b>	0.9710	0.9985	1.0000
Mod0	0.8879	0.8801	0.9966	0.9742	0.9710

Las Figuras II, IIa y IIb ilustran el impacto de los diferentes modelos sobre los estimadores de VA de la escuela. La dispersión de los estimadores que resultan de los modelos uno y cuatro muestran una tendencia general al acuerdo entre estimadores, con la excepción de la escuela mostrada en el triángulo de mayor tamaño rojo. La mayor diferencia entre la posición de la escuela en el rango dado por el modelo uno y su posición en el rango dado por el modelo cuatro es de 15 puestos (Figuras IIa y IIb).

FIGURA II. Dispersión de los estimadores de VA en el primer curso

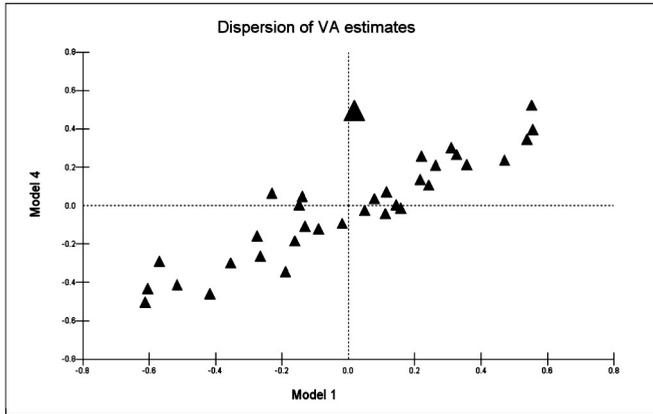
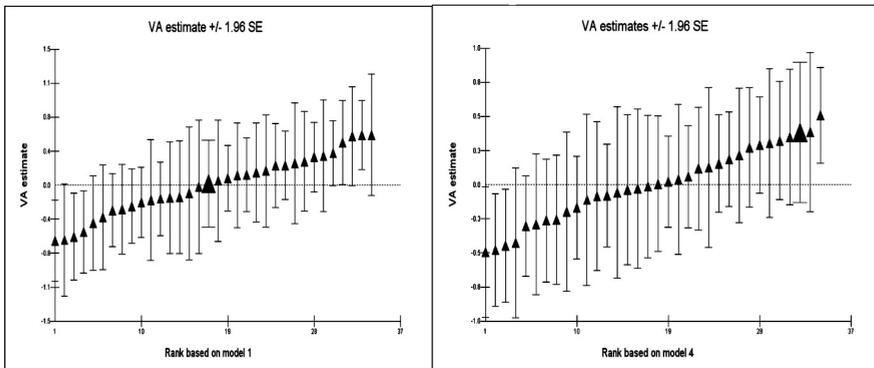
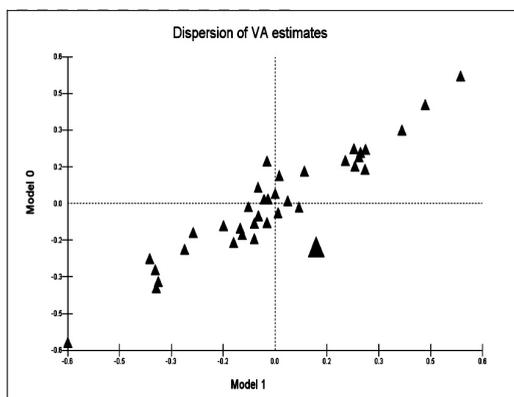


FIGURA IIa y IIb. Intervalos de Confianza (95%) para el VA



En el tercer curso, la correlación entre los estimadores de VA producidos por diferentes modelos es mayor que 0,93. El menor valor de la correlación entre rangos es de 0,88, y la mayor diferencia en las ordenaciones de los modelos es de 20 puestos.

FIGURA III. Dispersión de los estimadores de VA en el tercer curso

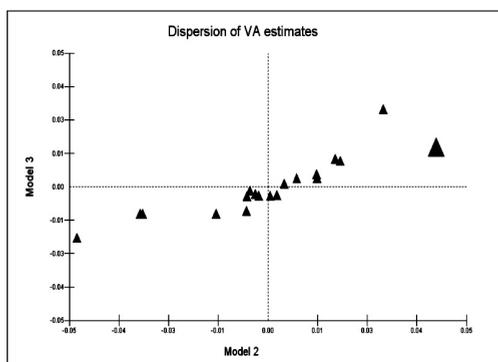


MATRIZ I. Correlación entre estimadores de VA – Quinto curso

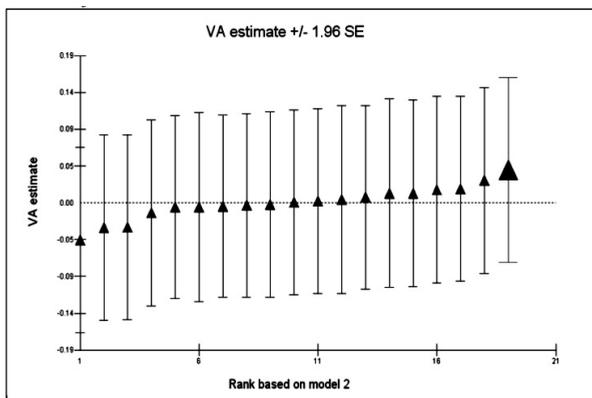
	Mod 3	Mod 2	Mod 1	Mod 0
Mod 3	1.0000			
Mod 2	0.6091	1.0000		
Mod 1	0.9980	0.6035	1.0000	
Mod 0	0.9155	0.8032	0.9199	1.0000

En el quinto curso la correlación entre los estimadores de VA generados por el modelo dos y los del modelo tres es de 0,61 (véase Figuras IV y V). La correlación en términos de la posición en los rangos es de 0,96.

FIGURA IV. Dispersión de los estimadores de VA en el quinto curso



FIGURAV. Intervalos de Confianza del 95% de estimadores de VA . Modelo dos



El modelo uno sugiere que en el quinto curso, después de controlar el rendimiento anterior y la educación de los padres, el VA no es estadísticamente diferente de cero.

En el séptimo curso, la correlación entre estimadores de VA producidos por los modelos de cero a cuatro es mayor que 0,96, y la correlación entre los respectivos rangos es mayor que 0,94. El rendimiento previo es el predictor más importante. En general, esta evidencia se mantiene también para el octavo curso, véase Figuras VII y VIII, que ilustran la comparación entre los estimadores de VA basados en el modelo cero (ajustando el rendimiento previo) y el modelo uno (que ajusta el rendimiento previo y la educación de los padres). Es importante tener en cuenta que el coeficiente de partición de la varianza (CPV) es bastante bajo en la educación elemental y en la secundaria inferior (véase Tabla III).

FIGURA VI. IC (VA)95%; modelo uno

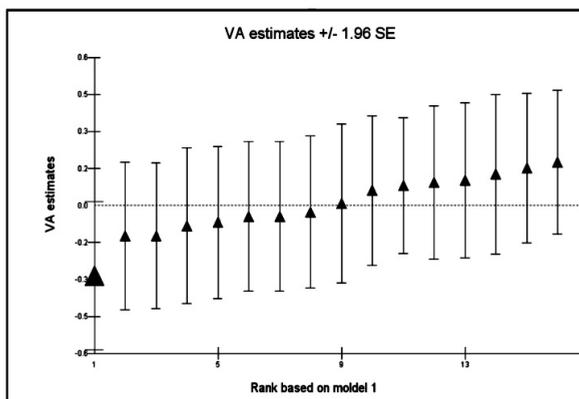


FIGURA VII. Dispersión de estimadores de VA en el octavo curso

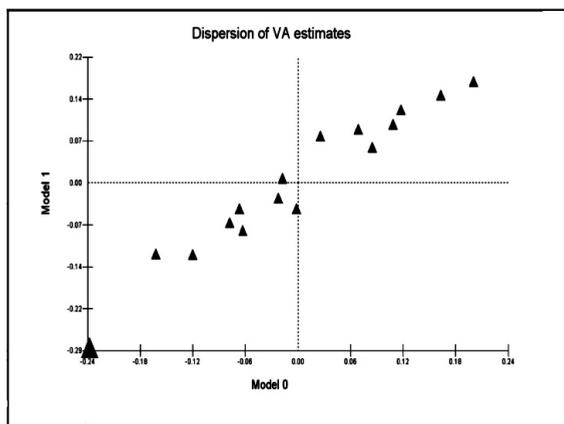


TABLA III. Estimadores de modelo nulo

	Grado				
	1 <sup>o</sup>	3 <sup>o</sup>	5 <sup>o</sup>	7 <sup>o</sup>	8 <sup>o</sup>
Intercepto	-0.019	-0.122	-0.005	-0.038	-0.003
	(0.092)	(0.088)	(0.086)	(0.091)	(0.100)
Parámetros aleatorios					
$\sigma_e^2$	0.155	0.164	0.077	0.080	0.096
	(0.067)	(0.066)	(0.045)	(0.048)	(0.057)
$\sigma_u^2$	0.848	0.857	0.918	0.911	0.897
	(0.073)	(0.072)	(0.079)	(0.080)	(0.086)
PVC	0.15	0.16	0.08	0.08	0.10

## Discusión

En este artículo hay dos resultados importantes. El primero es que el parámetro fijado ESE es estadísticamente significativo en todos los cursos, con la excepción del séptimo. El segundo es que el impacto de la elección de modelo (entendido como un conjunto diferente de variables de ajuste) sobre los estimadores de VA es particularmente importante en la Educación Primaria. En la educación elemental y en Secundaria inferior principalmente, los resultados de los modelos parecen confirmar los resultados de Ballou et al. (2004), sobre las relaciones entre ESE y rendimiento anterior. La

evidencia presentada parece sugerir que a lo largo de la trayectoria escolar, el rendimiento previo condensa los efectos del ESE. Es necesario investigar más sobre esto, pero si es cierto, esto constituye otro reto para la investigación futura sobre la equidad (buscar las escuelas que realmente «compensan» la desventaja en ESE).

Una de las limitaciones más importantes del trabajo presentado es la validez y fiabilidad de las variables de ESE, particularmente de la BCL. Períodos de observación más largos y datos poblacionales son también características muy relevantes para conocer el «verdadero» VA de la escuela. Continuamos investigando para complementar los análisis presentados con los datos longitudinales recogidos en 2006-07 y 2007-08.

Hay otros temas de interés relacionados a esta cuestión que también son objetivos para la investigación futura. Por ejemplo, Lockwood, McCaffrey, Hamilton, Stecher, Le y Martinez (2007) utilizaron datos longitudinales de una cohorte de estudiantes de escuela media de un distrito escolar y compararon varios modelos de VA (efectos del profesor). Encontraron que la variación intra profesores entre medidas de rendimiento es mayor que la variación entre profesores. Estos resultados sugieren que los modelos de VA son sensibles a las formas como se mide el rendimiento de los estudiantes, incluso cuando solamente hay una materia. En el número especial del *Journal of Educational and Behavioral Statistics* sobre modelos de VA, Reckase (2004) destaca la importancia de la evaluación: «los sofisticados procedimientos estadísticos descritos en estos artículos pueden proporcionar un final brillante a resultados de evaluación mal enfocados. Antes de otorgar mucha confianza a los resultados de estos análisis, es preciso investigar con mayor detalle el funcionamiento de las evaluaciones (p.120)».

## Referencias bibliográficas

- BALLOU, D., SANDERS, W. Y & WRIGHT, P. (2004). Controlling for Student Background in Value-Added Assessment of Teachers. *Journal of Educational and Behavioral Statistics*, 29, 37-66.
- BRAUN, H. (2005). *Using student progress to evaluate teachers: a primer on value-added models*. Policy Information Perspective, Princeton, NJ: ETS.
- BRAUN, H. Y & WAINER, H. (2007). Value-Added Modeling. In C.R. RAO Y & S. SINHARAY, S. (eds.) *Handbook of statistics 26, Psychometrics* (475-501). Amsterdam: Elsevier.
- COLEMAN, J. (1988). Social capital in the generation of human capital. *American Journal of Sociology*, 94, 95-120.

- FERRÃO, M.E. (2007). *Value-Added Models in Portugal*. Background paper prepared for the 1st meeting VAM project.
- FERRÃO, M.E. & GOLDSTEIN, H. (in press). Adjusting for Measurement Error in the Value Added Model: evidence from Portugal. *Quality & Quantity*.
- GRAY, J., JESSON, D., GOLDSTEIN, H., HEDGER, K. & RASBASH J. (1995). A multilevel analysis of school improvement: changes in schools' performances over time. *School Effectiveness and School Improvement*, 6, 97-114.
- LADD, H.F. & WALSH, R.P. (2002). Implementing value-added measures of school effectiveness: Getting the incentives right. *Economics of Education Review*, 21, 1-17.
- LITTLE, R.J.A. & RUBIN, D.B. (2002). *Statistical analysis with missing data*, (2nd edition) 9. Wiley: New York.
- LOCKWOOD, J.R., MCCAFFREY D.F., HAMILTON, L.S., STECHER, B.M., LE, V. & MARTINEZ, F. (2007). The sensitivity of value-added teacher effect estimates to different mathematics achievement measures. *Journal of Educational Measurement*, 44, 47-67.
- MCCAFFREY, D., LOCKWOOD, J.R., KORETZ, D.M. & HAMILTON, L.S. (2003). *Evaluating Value-Added Models for Teacher Accountability*. Santa Mónica, CA: RAND Corporation.
- RUBIN, D.B., STUART, E.A. & ZANUTTO, E.L. (2004). A potential outcomes view of value-added assessment in Education. *Journal of Educational and Behavioral Statistics*, 29, 103-116.
- SAMMONS, P., THOMAS, S. & MORTIMORE, P. (1997). *Forging links: effective schools and effective departments*. London: Chapman.
- THOMAS, S. Y & MORTIMORE, P. (1996). Comparison of value-added models for secondary school effectiveness., *Research Papers in Education*, 11, 5-33.
- VICENTE, P. (2006). O plano amostral do projecto 3EM. En M. E. FERRÃO, C. NUNES Y C. BRAUMANN (eds.), *Proceedings of the XIV Annual Conference of the Portuguese Statistical Society*: Lisboa: SPE.

**Dirección de contacto:** M<sup>a</sup> Eugénia Ferrão. Universidade da Beira Interior. Departamento de Matemática. Universidade da Beira Interior. Rua Marqués d'Avila e Bolama. 6200-001 Covilhã, Portugal. e-mail: mferrao@ubi.pt

## Anexo A

### Estimadores de los parámetros

Modelo 0	Curso 1		Curso 3		Curso 5		Curso 7		Curso 8	
Fixed	Estimador	ET								
cons	-0.015	0.090	-0.075	0.071	-0.022	0.048	-0.022	0.074	0.000	0.075
score_1	0.476	0.050	0.502	0.049	0.683	0.043	0.587	0.050	0.410	0.060
<b>Random</b>										
Level 2 var	0.172	0.065	0.089	0.041	0.011	0.014	0.052	0.032	0.035	0.032
Level 1 var	0.637	0.055	0.654	0.055	0.513	0.044	0.606	0.054	0.773	0.074
-2*log likelihood	745.828		788.950		636.784		650.076		606.510	

Modelo 1	Curso 1		Curso 3		Curso 5		Curso 7		Curso 8	
Fixed	Estimador	ET								
cons	-0.014	0.093	-0.046	0.076	-0.039	0.046	-0.042	0.077	-0.008	0.078
score_1	0.468	0.051	0.443	0.052	0.656	0.047	0.599	0.056	0.374	0.063
par_edu_stand	-0.113	0.054	-0.209	0.054	-0.117	0.047	-0.037	0.057	-0.113	0.065
<b>Random</b>										
Level 2 var	0.179	0.068	0.112	0.048	0.005	0.013	0.054	0.034	0.037	0.035
Level 1 var	0.636	0.057	0.632	0.055	0.479	0.043	0.642	0.059	0.770	0.079
-2*log likelihood	701.896		735.763		553.206		614.581		543.728	

Modelo 2	Curso 1		Curso 3		Curso 5		Curso 7		Curso 8	
Fixed	Estimador	ET								
cons	-0.017	0.096	-0.054	0.077	-0.039	0.046	-0.046	0.077	-0.006	0.078
score_1	0.468	0.052	0.447	0.053	0.653	0.047	0.600	0.056	0.370	0.064
par_edu	-0.116	0.057	-0.225	0.058	-0.106	0.050	-0.046	0.060	-0.104	0.070
par_edu_sch	0.023	0.167	0.101	0.140	-0.060	0.108	0.079	0.172	-0.069	0.189
<b>Random</b>										
Level 2 var	0.179	0.068	0.112	0.048	0.004	0.013	0.053	0.034	0.036	0.034
Level 1 var	0.636	0.057	0.632	0.055	0.479	0.043	0.642	0.059	0.771	0.079
-2*log likelihood	701.876		735.254		552.896		641.371		543.594	

Modelo 3	Curso 1		Curso 3		Curso 5		Curso 7		Curso 8	
Fixed	Estimador	ET								
cons	0.014	0.091	-0.070	0.074	0.071	0.058	-0.009	0.083	0.011	0.086
score_1	0.466	0.051	0.498	0.050	0.666	0.045	0.585	0.051	0.423	0.061
FSM	-0.172	0.131	-0.058	0.144	-0.235	0.092	-0.043	0.109	-0.027	0.131
<b>Random</b>										
Level 2 var	0.159	0.062	0.096	0.043	0.002	0.013	0.056	0.033	0.032	0.032
Level 1 var	0.637	0.055	0.657	0.055	0.513	0.046	0.608	0.054	0.788	0.077
-2*log likelihood	744.142		788.816		587.235		639.959		589.532	

Modelo 4	Curso 1		Curso 3		Curso 5		Curso 7		Curso 8	
Fixed	Estimador	ET								
cons	0.206	0.121	-0.076	0.101	0.206	0.095	0.180	0.204	0.069	0.159
score_1	0.467	0.050	0.498	0.050	0.653	0.045	0.586	0.051	0.420	0.062
FSM	-0.101	0.135	-0.061	0.149	-0.172	0.099	-0.024	0.110	-0.009	0.138
Prop_FSM	-1.117	0.503	0.039	0.561	-0.413	0.236	-0.616	0.613	-0.193	0.444
<b>Random</b>										
Level 2 var	0.130	0.054	0.100	0.044	0.000	0.000	0.052	0.032	0.031	0.027
Level 1 var	0.635	0.055	0.658	0.055	0.510	0.044	0.608	0.054	0.789	0.077
-2*log likelihood	739.456		788.856		584.232		638.969		589.344	

# Invarianza de la estructura de covarianzas de las medidas de rendimiento académico en estudios longitudinales en la transición de Educación Primaria a Secundaria

## Invariance of the covariance structure for the means of achievement variables in longitudinal studies dealing with the transition period from Primary Education to Secondary Education

José Luis Gaviria Soto, María Chantal Biencinto López y Enrique Navarro Asencio

*Universidad Complutense. Facultad de Educación. Departamento de Métodos de Investigación y Diagnóstico en Educación (MIDE). Madrid, España*

### Resumen

Dados los enormes cambios de carácter tanto psicobiológicos de los estudiantes como en la estructura organizativa y curricular del sistema educativo que experimentan en la transición entre Educación Primaria y Educación Secundaria, en el presente artículo se analiza la posibilidad de mantener la misma estructura longitudinal de base en los estudios de valor añadido con alumnos de Educación Primaria y con los de Educación Secundaria. Para ello, se estudia la estructura de varianza-covarianza de las variables de rendimiento en un estudio longitudinal realizado con 6.709 estudiantes agrupados en 153 centros educativos de la Comunidad de Madrid escolarizados entre 5º de Educación Primaria y 4º de ESO en los cursos académicos 2005-06 y 2006-07. Para el análisis de la invarianza de la estructura de varianza-covarianza del rendimiento en Matemáticas y Comprensión Lectora, se ha empleado, por un lado, el contraste mediante el estadístico M de Box y, por otro, el análisis de un modelo de ecuaciones estructurales de relaciones entre variables observadas, esto es, comenzar con un modelo que establezca la línea base de las relaciones en el tiempo. Los resultados obtenidos apoyan sólo parcialmente

la idea del mantenimiento de dicha estructura en las distintas cohortes o grupos, ya que existe una mayor semejanza en la estructura de Educación Primaria y primer ciclo de Educación Secundaria que entre los dos ciclos de secundaria. Esta variación asociada al tiempo se mantiene en los dos campos disciplinares que se han estudiado, Comprensión Lectora y Matemáticas, aunque cada uno de ellos tenga características diferenciadas.

*Palabras clave:* valor añadido, rendimiento, invarianza factorial, análisis de ecuaciones estructurales.

### **Abstract**

Given the huge changes that both the students and the structure of the educational system experience over the transition period from Primary Education to Secondary Education, this article studies the hypothesis of invariance of the same basic longitudinal structure on added value studies involving students from both Primary and Secondary Education levels. In order to do this, the covariance structure of achievement variables is studied in a longitudinal study carried out with 6,709 students grouped in 153 educational centres (Primary and Secondary Schools) in the Community of Madrid, from 5th level of Primary Education to 4th level of Secondary Education, in the academic years of 2005-06 and 2006-07. The invariance of the covariance structure of achievement variables in Mathematics and Reading literacy has been tested by means of, on one hand, the statistical M of Box and, on the other, the analysis of a structural equations model of relations between observed variables. Results obtained only partially support the idea of the maintenance of this structure in different cohorts or groups, since there is a stronger similarity between Primary Education and the first stage of Obligatory Secondary Education, than between the two stages of Obligatory Secondary Education. Reading literacy and Mathematics present their own characteristic temporal variation.

*Key Words:* Added value, achievement, factorial invariance, analysis of structural equations.

## **Introducción**

La transición de los alumnos de la etapa de Educación Primaria a la Educación Secundaria Obligatoria está acompañada por muchos cambios de carácter tanto psicobiológicos de los estudiantes como en la estructura organizativa y curricular del sistema educativo. Esto hace de ese paso un momento crítico que se manifiesta, entre otros indicadores, por un descenso del rendimiento escolar.

En esa misma transición, se produce también un cambio organizativo del proceso docente muy importante. Los alumnos pasan de tener un solo maestro que les acompaña a lo largo de casi toda la jornada escolar a tener a un grupo de distintos especialistas que se van turnando en la tarea de dar clases.

Tampoco es indiferente el hecho de que, conforme avanzan los alumnos en el sistema educativo, la complejidad de las materias estudiadas, a pesar de compartir título, aumenta significativamente. Ese aumento de complejidad curricular puede venir, y probablemente viene, acompañado de una mayor complejidad estructural de los factores latentes (Lizasoain y Joaristi, 2009).

Como consecuencia de ese aumento de la complejidad curricular, los modelos que sólo incorporan un factor como causa inmediata de las intercorrelaciones presentan en el ámbito de lo observado patrones correlacionales que difieren de los esperables en el nivel latente de las variables estructuradas (Blanco, González y Ordóñez, 2009).

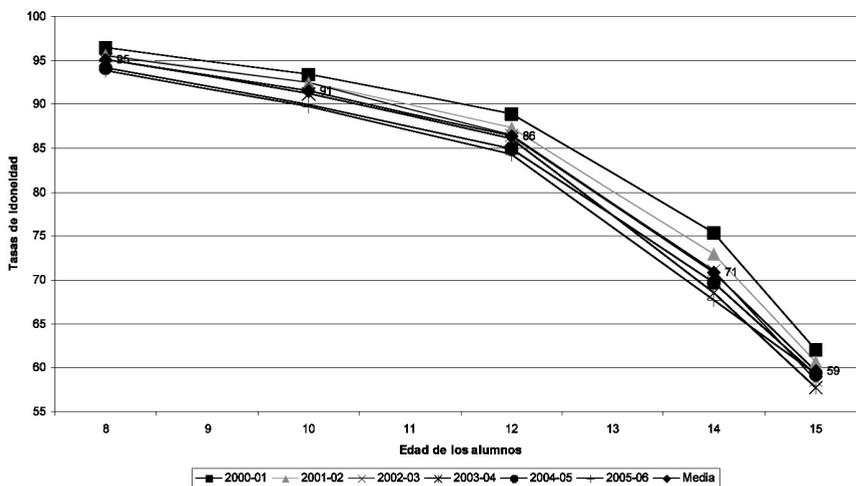
Por otra parte, la construcción de variables escaladas en una métrica común exige algunos diseños metodológicos y estructuras de las pruebas que faciliten la posterior equiparación y comparabilidad de los resultados.

Uno de estos indicadores sensibles a los problemas transicionales es la tasa de idoneidad. Su evolución refleja claramente la situación referida. La tasa de idoneidad se define como el porcentaje de alumnos que se encuentra matriculado en el curso que corresponde a su cohorte de edad (Instituto de Evaluación, 2006). Según las últimas estadísticas de la Educación en España disponibles (MEC, 2006), la tasa de idoneidad media para toda España a los 12 años se sitúa en un 84,3%. Los alumnos de la Comunidad de Madrid se sitúan ligeramente por encima de la media con una tasa bruta de 84,9%.

Sin embargo, la evolución de las tasas de idoneidad a partir de los 12 años muestra un importante descenso. El Gráfico I refleja la evolución de las tasas de idoneidad en cinco edades clave (desde los 8 a los 15 años) en la Comunidad de Madrid, que se corresponde con el momento de transición entre la primaria y la secundaria.

La investigación que se presenta en este artículo utiliza datos provenientes de un estudio longitudinal en el que tres cohortes de alumnos han sido medidos en cuatro ocasiones consecutivas, al inicio y al final de dos cursos académicos, en las dos materias instrumentales básicas de Matemáticas y Comprensión Lectora. El objetivo es tratar de comprobar si la estructura de varianzas covarianzas y de correlaciones entre las medidas de rendimiento es invariante en las tres cohortes analizadas.

GRÁFICO I. Tasas de idoneidad en la Comunidad de Madrid de 2001-02 a 2005-06



Fuente: elaboración propia a partir de las series de las estadísticas de enseñanza no universitaria del MEC

## Estudio empírico

### Objetivo del estudio

Este estudio pretende dar una respuesta a las siguientes preguntas:

- ¿La matriz de varianzas-covarianzas formada por las cuatro variables de respuesta del estudio longitudinal se mantiene estable entre las tres cohortes definidas, tanto para Comprensión Lectora como para Matemáticas?
- Si no es invariante, ¿entre qué grupos existe una mayor diferencia?
- Como consecuencia de lo anterior, ¿es posible utilizar el mismo modelo de valor añadido para primaria y secundaria o, por el contrario, se necesitan modelos diferentes o, incluso, un modelo distinto para cada cohorte?

### Muestra y población. Estructura de las cohortes

La población objeto de estudio fue el conjunto de alumnos (tanto de centros públicos, como privados y privados concertados) de la Comunidad de Madrid escolarizados en el

último ciclo de Educación Primaria y en Educación Secundaria Obligatoria, que asciende a un total de 343.746 repartidos en 1.798 escuelas en el curso académico 2005-06. El dominio específico de este estudio es la Comunidad Autónoma de Madrid, constituida por toda la red de centros públicos, privados concertados y privados, que componen subdominios específicos dentro de la misma. En este dominio, se definieron cinco estratos correspondientes con las cinco áreas territoriales (Capital, Norte, Sur, Este y Oeste) de la Comunidad.

Para el cálculo inicial del tamaño necesario en cada una de las cohortes, se utilizó un muestreo aleatorio simple utilizando el tamaño de la población de referencia, un nivel de confianza del 95% y un error muestral máximo de  $0,1\sigma$ . La segunda parte del proceso de consistió en un muestreo por conglomerados, éste tiene un efecto negativo, es decir, el tamaño de la muestra aumentará, con respecto al muestreo aleatorio simple, en función del tamaño del conglomerado y de la autocorrelación.

Para llevar a cabo esta corrección, se tomaron los valores de 25 como tamaño medio del conglomerado y una correlación intraclase de 0,2 (Kiplinger, 2004; Zvoch y Stevens, 2003; Navarro y Redondo, 2007). Finalmente, se introdujo la modificación debida a la mortalidad experimental que se produce en este tipo de estudios longitudinales.

La muestra final está compuesta por 174 centros repartidos en tres grados, seleccionados con probabilidad proporcional a su tamaño y estratificada por la titularidad y el área territorial. El total de alumnos implicados llegó hasta 13.000, medidos en cuatro ocasiones a lo largo de los cursos académicos 2005-06 y 2006-07.

Para llevar a cabo este estudio, se han utilizado sólo los datos de los sujetos con puntuaciones en las cuatro aplicaciones; es decir, los de aquéllos de los que se tenía registro completo porque participaron en las cuatro evaluaciones llevadas a cabo. La muestra definitiva queda configurada como se muestra en la Tabla I.

**TABLA I.** Configuración definitiva de las muestras

	MATEMÁTICAS		COMPRESIÓN LECTORA	
	ALUMNOS	CENTROS	ALUMNOS	CENTROS
Cohorte 1	2751	92	2739	109
Cohorte 2	2152	61	2128	65
Cohorte 3	1806	59	1887	

## Características del diseño

El diseño longitudinal para esta investigación supone la incorporación de cuatro ocasiones de medida a lo largo de dos cursos académicos. De esta forma, se mide a los

alumnos al inicio y al final del curso académico, durante dos años consecutivos (2005-06 y 2006-07). Se definieron tres segmentos educativos de interés: el último ciclo de Educación Primaria y los dos ciclos de Educación Secundaria Obligatoria.

Cada uno de estos segmentos educativos constituye una cohorte. Para los alumnos que en 2005-06 comenzaron en 5º de primaria (cohorte 1) se han recogido datos en octubre de 2005 (primera medición) y junio de 2006 (segunda medición). En el curso 2006-07, cuando estos alumnos se encontraban en 6º de primaria, se recogió información en noviembre de 2006 (tercera medición) y junio de 2007 (cuarta medición). Con las dos cohortes restantes correspondientes al primer ciclo de secundaria (cohorte 2) y segundo ciclo de secundaria (cohorte 3), se ha procedido de manera análoga. La Tabla II muestra la estructura de la recogida de datos longitudinal asociada a esta investigación.

TABLA II. Estructura de recogida longitudinal de datos

Cohorte	Curso	2005-06									2006-07											
		2005			2006						2006			2007								
1	5º EP	Oct	Nov	Dic	Ene	Feb	Marz	Abr	May	Jun												
	6º EP										Jul	Agos	Sept	Oct	Nov	Dic	Ene	Feb	Marz	Abr	May	Jun
2	1ºESO	Oct	Nov	Dic	Ene	Feb	Marz	Abr	May	Jun												
	2ºESO										Jul	Agos	Sept	Oct	Nov	Dic	Ene	Feb	Marz	Abr	May	Jun
3	3ºESO	Oct	Nov	Dic	Ene	Feb	Marz	Abr	May	Jun												
	4ºESO										Jul	Agos	Sept	Oct	Nov	Dic	Ene	Feb	Marz	Abr	May	Jun

Este tipo de diseño de recogida de información longitudinal nos permite obtener medidas del rendimiento de los alumnos al final de tres de las etapas clave del Sistema Educativo Español (final de la Educación Primaria y de los dos ciclos de Educación Secundaria Obligatoria).

Al mismo tiempo se evitó el grave problema que para el diseño de investigación supone que una cohorte cambie de ciclo durante el período de la recogida de datos. En efecto, los alumnos que comenzaron en 5º de Educación Primaria se encontraban al final de 6º de Educación Primaria en el mes de junio de 2007 cuando se recogieron los últimos datos de cada cohorte. Si se hubiese comenzado tomando datos a alumnos de 6º de Educación Primaria, en el segundo año del estudio estos alumnos habrían cambiado de etapa y, en muchos casos, de centro. La consecuencia de esto último hubiera sido una gran pérdida de datos longitudinales, dado que muchos alumnos habrían cambiado a otros centros no incluidos en la muestra estudiada.

Esta ventaja señalada implica también una importante limitación para este tipo de estudios, ya que, aunque tenemos datos de una cohorte de Educación Primaria y dos

de Educación Secundaria, no disponemos de datos de una cohorte en la que se incluya la transición de una etapa a la otra. No se escapará al lector la dificultad de recabar este tipo de información.

Esto hace todavía más importante la necesidad de comprobar la constancia de la estructura de los datos analizados en las etapas del Sistema Educativo.

## Procedimientos y modelos de análisis

Para contrastar las hipótesis planteadas se utilizaron dos procedimientos. En el primero, se utilizó la prueba M de Box para comprobar la igualdad de las matrices de covarianzas entre cohortes. El estadístico M puede aproximarse por una distribución F de Snedecor si se cumplen los supuestos del mismo, el más importante de los cuales es el de normalidad de las distribuciones implicadas. El cumplimiento de este supuesto se estudió mediante las pruebas de Kolmogorov-Smirnov y Shapiro y Wilks.

El segundo procedimiento consistió en el contraste de un modelo de ecuaciones estructurales multigrupo.

La estructura longitudinal del estudio impone analizar el modelo compuesto por las cuatro variables extraídas de las puntuaciones obtenidas por los alumnos en las cuatro aplicaciones, tanto en Comprensión Lectora como en Matemáticas. El estudio de la variación de esta estructura entre cohortes permite comprobar si el cambio que se refleja en algunos indicadores en los alumnos en la transición de primaria a secundaria supone también variaciones en los patrones de correlaciones entre las medidas consecutivas tan importante que impliquen modelos de valor añadido distintos.

La prueba M de Box (Box, 1949) contrasta la hipótesis nula de igualdad de las matrices de varianzas-covarianzas poblacionales.

El estadístico M de Box toma la siguiente forma:

$$M = (n - g) \log |S| - \sum_{j=1}^g (n_j - 1) \log |S_j| \quad (1)$$

Donde  $S$  es la matriz de varianzas-covarianzas combinada y  $S_j$  la de cada grupo  $j$ ,  $n$  es el número total de casos,  $n_j$  hace referencia al número de casos de cada grupo  $j$  y  $g$  es el número de grupos. El estadístico M no tiene distribución muestral conocida; no obstante, puede aproximarse por un estadístico con distribución F de Snedecor. Por tanto, un valor de dicha F significativo indicará que la estructura de covarianzas no es invariante entre cohortes.

El test M de Box es sensible a la falta de normalidad multivariante, es decir, matrices iguales pueden aparecer como significativamente diferentes si no existe normalidad (Pardo y Ruiz, 2002).

Esta especial sensibilidad a las violaciones de los supuestos de normalidad y dado el importante tamaño de la muestra con que se contaba, que hacía más probable el rechazo de la hipótesis de nulidad, hizo aconsejable complementar el estudio con otra estrategia de análisis algo más robusta y con índices de ajuste más adecuados a la naturaleza de los datos manejados.

Esta estrategia es el análisis de modelos de ecuaciones estructurales multigrupo (Arbuckle, 2006). Este tipo de análisis permite imponer restricciones sucesivas al modelo planteado; en nuestro caso, como muestra la Figura III, sólo consta de varianzas y covarianzas entre puntuaciones; por tanto, la única restricción es la fijación de dicha matriz.

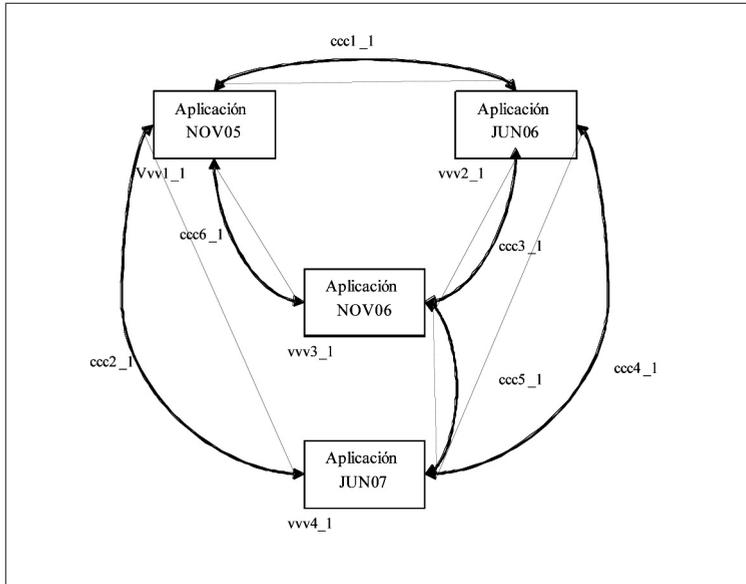
Es importante señalar en este punto que nuestro objetivo no es probar un modelo de rendimiento académico, ya que éste sería mucho más complejo e implicaría la introducción de distintos predictores con el objetivo de maximizar la varianza explicada. Como ya se ha señalado anteriormente, un estudio longitudinal supone el comenzar con un modelo que establece la línea base, es decir, cómo es la relación de las medidas dependientes con el tiempo o, dicho de otro modo, de qué manera se produce la sucesión de las variables dependientes cuyo progreso habrá que explicar con los predictores adecuados.

El modelo sometido a prueba por tanto es el que aparece en la Figura I, con la restricción de que las correlaciones permanecen invariantes en las tres cohortes.

Para el contraste de este modelo, se utilizó el programa AMOS 7.0 (Arbuckle, 2006). Para llevar a cabo el proceso de análisis, se han comparado diferentes modelos que varían en las restricciones impuestas. El modelo saturado asume que todos los parámetros posibles son significativamente distintos de cero; se trata de un modelo de ajuste perfecto y que, por tanto, refleja completamente toda la información de los datos. En el otro extremo, se encuentra el modelo fijo; éste asume que todos los parámetros son iguales entre cohortes y, por tanto, un buen ajuste de este modelo indicaría la semejanza entre los diferentes grupos estudiados. Entre estos dos modelos extremos, se sitúan otros que permiten establecer comparaciones entre pares de cohortes.

Todas las estimaciones se han realizado por el procedimiento de máxima verosimilitud, tal y como está implementado en el mencionado programa AMOS. Para evaluar el ajuste del modelo, se han utilizado, aparte del estadístico de bondad de ajuste  $\chi^2$ , otros índices que juzgan la adecuación del modelo (Bollen y Long, 1993), ya que, en ocasiones, valores altos de  $\chi^2$  pueden producirse por el incumplimiento de los

FIGURA I. Modelo de contraste



supuestos de normalidad (Batista y Coenders, 2000; Baldwin y Caldwell, 2003; Schumacker y Lomax, 2004). Se ha utilizado para tal caso el índice de bondad de ajuste GFI (Jöreskog y Sörbom, 1989) y su ajustado AGFI. Valores cercanos a uno en estos índices suponen un buen ajuste. Por encima de 0,90 son aceptables. También se ha analizado la raíz media cuadrática del error de estimación (RMSEA) (Browne y Cudeck, 1993) y su intervalo de confianza. Valores inferiores a 0,05 señalan un correcto ajuste de los datos. Algunos autores consideran que valores entre 0,05 y 0,08 también son aceptables (Hair, Anderson, Tatham y Black, 1999).

## Resultados

Comenzamos presentando los resultados correspondientes a Comprensión Lectora y seguiremos con los de Matemáticas. Se presentan, en primer lugar, los resultados de la prueba M de Box, seguidamente la prueba de normalidad de las variables de respuesta y, por último, los del análisis multigrupo.

## Resultados en Comprensión Lectora

En la Tabla III, podemos ver la matriz de varianzas-covarianzas entre las medidas de Comprensión Lectora. En dicha matriz, podemos observar cómo los patrones que se producen en una cohorte se asemejan mucho a los que se producen en cualquiera de las otras dos.

TABLA III. Matriz de Varianzas - Covarianzas

Cohorte 1		CLNOV05	CLJUN06	CLNOV06	CLJUN07
	CLNOV05	,669	,385	,394	,372
CLJUN06		,657	,401	,383	
CLNOV06			,691	,415	
CLJUN07				,695	
Cohorte 2		CLNOV05	CLJUN06	CLNOV06	CLJUN07
	CLNOV05	,701	,389	,406	,373
	CLJUN06		,734	,475	,409
	CLNOV06			,837	,465
CLJUN07				,752	
Cohorte 3		CLNOV05	CLJUN06	CLNOV06	CLJUN07
	CLNOV05	,642	,323	,292	,321
	CLJUN06		,638	,328	,329
	CLNOV06			,527	,358
CLJUN07				,731	

Si nos fijamos en las diferencias en las varianzas y covarianzas de las distintas aplicaciones, vemos que los patrones en ningún caso reflejan una mayor distancia entre la cohorte de primaria y secundaria que entre los dos ciclos de secundaria. Como se puede observar en la Tabla IV, existe mayor diferencia entre el primer ciclo de secundaria y el segundo, es decir, entre la cohorte 2 y 3, que entre la cohorte de primaria y los dos ciclos de secundaria. Por ejemplo, la diferencia en la covarianza de la puntuación CLNOV05 y CLJUN06 es de 0,004 entre la cohorte de primaria y la de primer ciclo de secundaria; en cambio, esta misma diferencia de covarianzas tienen un valor de 0,066 entre los dos ciclos de secundaria. En cualquier caso es evidente que se trata de mínimas diferencias.

Por otra parte, observando la matriz de correlaciones (Tabla V) podemos comprobar la intensa relación existente entre las aplicaciones en cada una de las tres cohortes. Prácticamente la totalidad de las correlaciones están por encima de 0,50 y resultan significativas al 99%.

Como cabía esperar, las correlaciones entre dos aplicaciones sucesivas presentan siempre el mayor valor. Conforme transcurre más tiempo entre aplicaciones el valor de la correlación se amortigua. Este fenómeno se produce de igual modo en las tres

TABLA IV. Diferencias en las varianzas y covarianzas entre cohortes

Diferencia 1-2		CLNOV05	CLJUN06	CLNOV06	CLJUN07
	CLNOV05		-0,032	-0,004	-0,012
CLJUN06			-0,077	-0,074	-0,026
CLNOV06				-0,146	-0,05
CLJUN07					-0,057
Diferencia 1-3		CLNOV05	CLJUN06	CLNOV06	CLJUN07
	CLNOV05		0,027	0,102	0,051
	CLJUN06		0,019	0,073	0,054
	CLNOV06			0,164	0,057
CLJUN07					-0,036
Diferencia 2-3	2	CLNOV05	CLJUN06	CLNOV06	CLJUN07
	CLNOV05		0,059	0,114	0,052
	CLJUN06		0,096	0,147	0,08
	CLNOV06			0,31	0,107
	CLJUN07				0,021

cohortes analizadas. Es interesante comprobar que los valores absolutos de las correlaciones disminuyen proporcionalmente a la edad media de la cohorte. En efecto, las correlaciones de la cohorte 3 son en términos absolutos más bajas que las de la 2 y las de la 1. Pero en los tres casos se reproduce un patrón similar de amortiguación.

TABLA V. Matriz de Correlaciones

Cohorte 1		CLNOV05	CLJUN06	CLNOV06	CLJUN07
	CLNOV05		1	,581 <sup>(***)</sup>	,579 <sup>(***)</sup>
CLJUN06			1	,594 <sup>(***)</sup>	,566 <sup>(**)</sup>
CLNOV06				1	,598 <sup>(**)</sup>
CLJUN07					1
Cohorte 2		CLNOV05	CLJUN06	CLNOV06	CLJUN07
	CLNOV05		,542 <sup>(***)</sup>	,530 <sup>(***)</sup>	,513 <sup>(**)</sup>
	CLJUN06		1	,606 <sup>(***)</sup>	,551 <sup>(**)</sup>
	CLNOV06			1	,586 <sup>(**)</sup>
CLJUN07					1
Cohorte 3		CLNOV05	CLJUN06	CLNOV06	CLJUN07
	CLNOV05		,505 <sup>(***)</sup>	,502 <sup>(***)</sup>	,468 <sup>(**)</sup>
	CLJUN06		1	,566 <sup>(***)</sup>	,481 <sup>(**)</sup>
	CLNOV06			1	,577 <sup>(**)</sup>
CLJUN07					1

\*\* La correlación es significativa al nivel 0,01 (bilateral).

En la Tabla VI, aparecen los resultados de la prueba M de Box. Como ya hemos comentado en párrafos anteriores, el estadístico de Box contrasta la hipótesis nula de igualdad entre las matrices de covarianzas poblacionales.

Los resultados de la prueba M de Box reflejan, inicialmente, diferencias significativas entre las matrices de covarianzas de las diferentes cohortes. Dado el tamaño muestral con el que se ha realizado el estudio, resulta interesante comprobar hasta qué punto distintas muestras aleatorias con tamaño menor reflejan esos mismo resultados. Para ilustrar esto se seleccionaron submuestras aleatorias de 3.000, 1.000, 300 y 100 sujetos.

Con una selección aleatoria de 100 sujetos, la prueba deja de mostrar diferencias al nivel de 0,01 bajando la significatividad hasta 0,045. Esto es debido a la sensibilidad de esta prueba al tamaño de la muestra. En la Tabla VI, podemos ver la evolución del estadístico y su probabilidad asociada en función de la reducción muestral llevada a cabo.

**TABLA VI.** Resultados de la prueba M de Box

M de Box	Muestra completa (6754 sujetos)	3000 sujetos	1000 sujetos	300 sujetos	100 sujetos
	186,029	120,644	45,229	42,118	34,067
<b>F</b> Aprox.	9,293	6,020	2,248	2,063	1,594
<b>Sig.</b>	,000	,000	,001	,003	,045

A la vista de estos resultados, cabe preguntarse si la violación de los supuestos de la prueba puede influir en los valores obtenidos. La prueba de Kolmogorov-Smirnov y la de Shapiro y Wilk nos indican hasta qué punto nuestros datos son consistentes con la hipótesis de normalidad multivariante. En el caso de Comprensión Lectora, las cuatro variables carecen de normalidad en cada una de las tres cohortes estudiadas.

La segunda parte del proceso de análisis consiste, como ya hemos comentado en el epígrafe anterior, en poner a prueba el modelo de interrelación de las matrices de covarianzas entre las cohortes mediante un análisis multigrupo en el contexto de análisis de ecuaciones estructurales. El resumen de los resultados más relevantes se presenta en la Tabla IX. Los modelos comparados en dicha tabla son los siguientes:

**TABLA VIII.** Modelos de comparación entre cohortes

MODELO	DEFINICIÓN	PARÁMETROS ESTIMADOS
<b>Saturado</b>	Todos los parámetros libres	4x3 varianzas + 6x3 covarianzas= 30
<b>1</b>	Con covarianzas fijas entre cohortes 1 y 2	4x3 varianzas + 6x2 covarianzas (las de las cohortes 1 y 3; recordemos que las de la cohorte 2 son iguales a las de la 1)=24
<b>2</b>	Con covarianzas fijas entre cohortes 1 y 3	4x3 varianzas + 6x2 covarianzas (las de las cohortes 1 y 2; recordemos que las de la cohorte 3 son iguales a las de la 1)=24
<b>3</b>	Con covarianzas fijas entre cohortes 2 y 3	4x3 varianzas + 6x2 covarianzas (las de las cohortes 1 y 2; recordemos que las de la cohorte 3 son iguales a las de la 2)=24
<b>Fijo</b>	Todos los parámetros fijos	4 varianzas + 6 covarianzas=10

Los valores de  $\chi^2$  correspondientes a estos modelos no indican un buen ajuste de los mismos, siendo valores altos debido muy probablemente a la no-normalidad de las variables y al gran tamaño de la muestra. En cambio, el resto de índices muestran unos valores que indican en términos generales que estamos trabajando con modelos muy plausibles. El modelo con todos los parámetros fijos, es decir, el que nos mostraría la igualdad en la matriz de varianzas-covarianza, con valores de GFI y AGFI de 0,987 y 0,980, respectivamente, y con un valor de RMSEA de 0,035, también su intervalo de confianza se sitúa por debajo de 0,05, señala en esa dirección. Podemos concluir, por tanto, que la estructura de las variables de respuesta en comprensión lectora es semejante en primaria y secundaria.

Si nos centramos ahora en las comparaciones por pares de cohortes (Tabla IX), podemos destacar la mayor similitud existente entre primaria y primer ciclo de secundaria (Modelo 1), con valores de ajuste muy buenos, que entre los dos ciclos de Educación Secundaria (Modelo 3).

Por tanto, la calidad y ajuste de los diferentes modelos son muy satisfactorios, como se puede comprobar en la Tabla IX.

TABLA IX. Resumen de la estimación de los modelos de comparación entre cohortes para Comprensión Lectora

Modelos	NPAR	CMIN	DF	P	CMIN/DF	GFI	AGFI	RMSEA	LO 90	HI 90
Modelo saturado	30	,000	0			1,000				
Modelo 1	24	17,699	6	,007	2,950	,999	,993	,017	,008	,026
Modelo 2	24	25,874	6	,000	4,312	,998	,991	,022	,014	,031
Modelo 3	24	42,332	6	,000	7,055	,997	,984	,030	,022	,039
Modelo fijo	10	185,996	20	,000	9,300	,987	,980	,035	,031	,040

Estos resultados nos llevan a pensar que la posible discontinuidad se produce no tanto en el inicio de la Educación Secundaria, sino más bien al avanzar en esta última. Es posible que las causas de este salto se presenten desde un principio en el cambio entre Educación Primaria y Educación Secundaria, pero sus efectos se hacen especialmente visibles al acercarse el último tramo de la Educación Secundaria.

Como consecuencia de este resultado, podemos preguntarnos si el cambio implica una variación en la importancia que la primera aplicación de cada cohorte tiene en la predicción de los rendimientos posteriores. Para tratar de responder a esta cuestión, se sometieron a contraste tres modelos para cada materia. En ellos, se comparan los ajustes cuando las correlaciones de la primera aplicación con las demás se mantienen constantes en las sucesivas cohortes, en un patrón de comparaciones similar al anterior.

Para llevar a cabo este paso, se ha seguido el mismo procedimiento: primero la estimación del modelo saturado y del modelo fijo (en este caso, es el que mantiene fijas las covarianzas sólo de la primera aplicación con el resto entre cohortes) y, a continuación, la estimación de diferentes modelos con restricciones entre pares de cohortes. Los modelos estimados son los siguientes:

TABLA X. Modelos

MODELO	DEFINICIÓN	PARÁMETROS ESTIMADOS
<b>Saturado</b>	Todos los parámetros libres	4x3 varianzas + 6x3 covarianzas= 30
<b>1</b>	Con las covarianzas de la primera aplicación con el resto fijas, entre la cohortes 1 y 2	4x3 varianzas + 3x3 covarianzas (son las cov. entre las 3 últimas aplicaciones en las tres cohortes) + 3x2 covarianzas (las de la primera aplicación con el resto en la cohorte 1 y 3; recordemos que las de la cohorte 2 son las mismas que las de la 1, las fijas)=27
<b>2</b>	Con las covarianzas de la primera aplicación con el resto fijas, entre la cohortes 1 y 3	4x3 varianzas + 3x3 covarianzas (son las cov. entre las 3 últimas aplicaciones en las tres cohortes) + 3x2 covarianzas (las de la primera aplicación con el resto en la cohorte 1 y 2; recordemos que las de la cohorte 3 son las mismas que las de la 1, las fijas)=27
<b>3</b>	Con las covarianzas de la primera aplicación con el resto fijas, entre la cohortes 2 y 3	4x3 varianzas + 3x3 covarianzas (son las cov. entre las 3 últimas aplicaciones en las tres cohortes) + 3x2 covarianzas (las de la primera aplicación con el resto en la cohorte 1 y 2; recordemos que las de la cohorte 3 son las mismas que las de la 2, las fijas)=27
<b>Fijo</b>	Con las covarianzas de la primera aplicación con el resto fijas	4x3 varianzas + 3x3 covarianzas (son las cov. entre las 3 últimas aplicaciones en las tres cohortes) + 3 covarianzas (las de la primera aplicación con el resto en una cohorte) =24

El modelo fijo muestra unos buenos índices de ajuste (Tabla XI) con un valor de GFI de 0,999 y un RMSEA de 0,017. Este modelo sólo es superado por el que mantiene fijas las covarianzas de la primera aplicación entre las cohortes 1 y 2, con un ajuste casi perfecto.

TABLA XI. Importancia de la primera aplicación como predictor (Comprensión Lectora)

Modelos	NPAR	CMIN	DF	P	CMIN/DF	GFI	AGFI	RMSEA	LO 90	HI 90
Modelo saturado	30	,000	0			1,000				
Modelo 1	27	,465	3	,927	,155	1,000	1,000	,000	,000	,006
Modelo 2	27	24,293	3	,000	8,098	,998	,982	,032	,021	,045
Modelo 3	27	24,960	3	,000	8,320	,998	,982	,033	,022	,045
Modelo fijo	24	17,699	6	,007	2,950	,999	,993	,017	,008	,026

Podemos afirmar que la primera aplicación tiene la misma importancia como predictor en las tres cohortes, encontrando mayores diferencias entre los dos ciclos de secundaria que entre cualquiera de éstos con la cohorte de primaria.

## Resultados en Matemáticas

La Tabla XII permite observar las varianzas y covarianzas de las variables de respuesta en las tres cohortes.

TABLA XII. Matriz de Varianzas - Covarianzas

Cohorte 1		MATNOV05	MATJUN06	MATNOV06	MATJUN07
	MATNOV05	0,776	0,547	0,358	0,337
	MATJUN06		0,811	0,412	0,391
	MATNOV06			0,358	0,261
	MATJUN07				0,331
Cohorte 2		MATNOV05	MATJUN06	MATNOV06	MATJUN07
	MATNOV05	0,733	0,524	0,392	0,347
	MATJUN06		0,739	0,42	0,388
	MATNOV06			0,424	0,306
	MATJUN07				0,406
Cohorte 3		MATNOV05	MATJUN06	MATNOV06	MATJUN07
	MATNOV05	0,735	0,418	0,322	0,292
	MATJUN06		0,802	0,441	0,381
	MATNOV06			0,534	0,326
	MATJUN07				0,421

De la misma forma que ocurría en Comprensión Lectora, las diferencias brutas reflejan una mayor discrepancia entre las covarianzas de los dos ciclos de secundaria que entre las cohortes de primaria y primer ciclo de secundaria (tabla 13). Podemos tomar como ejemplo el valor de la diferencia entre la covarianza de la primera aplicación (MATNOV05) y la última toma de datos (MATJUN07). Esta última tiene un valor absoluto de 0,01 entre la cohorte de primaria y primer ciclo de secundaria; llega a la cantidad de 0,045 entre la muestra de primaria y la de segundo ciclo de secundaria para finalmente ascender a 0,055 entre las dos cohortes de secundaria (Tabla XIII).

TABLA XIII. Diferencias en las varianzas y covarianzas entre cohortes

Diferencia 1-2		MATNOV05	MATJUN06	MATNOV06	MATJUN07
	MATNOV05	0,043	0,023	-0,034	-0,01
	MATJUN06		0,072	-0,008	0,003
	MATNOV06			-0,066	-0,045
	MATJUN07				-0,075
Diferencia 1-3		MATNOV05	MATJUN06	MATNOV06	MATJUN07
	MATNOV05	0,041	0,129	0,036	0,045
	MATJUN06		0,009	-0,029	0,01
	MATNOV06			-0,176	-0,065
	MATJUN07				-0,09
Diferencia 2-3		MATNOV05	MATJUN06	MATNOV06	MATJUN07
	MATNOV05	-0,002	0,106	0,07	0,055
	MATJUN06		-0,063	-0,021	0,007
	MATNOV06			-0,11	-0,02
	MATJUN07				-0,015

Los valores de las correlaciones entre las puntuaciones de las cuatro aplicaciones en Matemáticas son superiores, en las tres cohortes, a los obtenidos en la materia de Comprensión Lectora. Recordemos que en Comprensión Lectora las correlaciones variaban en torno a 0,5 y, en este caso, en las dos primeras cohortes llegan a 0,765 entre la segunda y tercera aplicación (Tabla XIV). En la tercera cohorte, los valores son ligeramente más bajos; no obstante, son correlaciones significativas.

TABLA XIV. Matriz de Correlaciones

Cohorte 1		MATNOV05	MATJUN06	MATNOV06	MATJUN07
	MATNOV05	1	,689 <sup>(***)</sup>	,679 <sup>(***)</sup>	,665 <sup>(***)</sup>
MATJUN06			1	,765 <sup>(***)</sup>	,754 <sup>(***)</sup>
MATNOV06				1	,758 <sup>(***)</sup>
MATJUN07					1
Cohorte 2		MATNOV05	MATJUN06	MATNOV06	MATJUN07
	MATNOV05	1	,712 <sup>(***)</sup>	,704 <sup>(***)</sup>	,637 <sup>(***)</sup>
MATJUN06			1	,750 <sup>(***)</sup>	,709 <sup>(***)</sup>
MATNOV06				1	,737 <sup>(***)</sup>
MATJUN07					1
Cohorte 3		MATNOV05	MATJUN06	MATNOV06	MATJUN07
	MATNOV05	1	,544 <sup>(***)</sup>	,514 <sup>(***)</sup>	,525 <sup>(***)</sup>
MATJUN06			1	,674 <sup>(***)</sup>	,656 <sup>(***)</sup>
MATNOV06				1	,688 <sup>(***)</sup>
MATJUN07					1

\*\* La correlación es significativa al nivel 0,01 (bilateral)

La prueba M de Box muestra cómo, de forma similar a la materia de Comprensión Lectora, la matriz de varianzas-covarianzas no es invariante entre las cohortes cuando se utiliza toda la muestra. No obstante, con la progresiva reducción muestral, disminuye de modo muy importante el valor de F (ver Tabla XV).

TABLA XV. Resultados de la prueba M de Box

M de Box	Muestra completa (6689 sujetos)	3000 sujetos	1000 sujetos	300 sujetos	100 sujetos
	668,962	307,757	152,349	86,549	36,821
F Aprox.	33,418	15,357	7,572	4,240	1,728
Sig.	,000	,000	,000	,000	,023

A la luz de estos resultados, es conveniente analizar, de la misma forma que hicimos en Comprensión Lectora, la normalidad de las cuatro variables que componen la estructura longitudinal del estudio. Los datos de las dos pruebas realizadas reflejan la falta de normalidad en la distribución de las variables. No obstante, podemos excluir

de esta tendencia a la variable correspondiente a la tercera aplicación (MATNOV06) en la cohorte de primer ciclo de Educación Secundaria Obligatoria (cohorte 2).

En la segunda fase del análisis, tenemos los resultados obtenidos a través del modelo de ecuaciones estructurales en su versión multigrupo. El resumen de la información podemos encontrarlo en la Tabla XVI. No es necesario volver a definir cada modelo porque son iguales a los estimados con Comprensión Lectora.

Los índices de ajuste empeoran ligeramente si los comparamos con los obtenidos en Comprensión Lectora. Los valores de GFI y AGFI son de 0,949 y 0,924, respectivamente, para el modelo fijo; en cambio, el valor de RMSEA con un valor de 0,07 no refleja un buen ajuste.

De la misma forma que ocurría en Comprensión Lectora, el modelo que mejor ajusta es aquél que mantiene fijas las covarianzas entre la primera y segunda cohorte, es decir, entre Educación Primaria y primer ciclo de Educación Secundaria Obligatoria. Las cohortes que muestran mayores diferencias son la uno y la tres como indican los resultados del modelo 2.

TABLA XVI. Resumen de la estimación de los modelos de comparación entre cohortes para Matemáticas

Modelos	NPAR	CMIN	DF	P	CMIN/DF	GFI	AGFI	RMSEA	LO 90	HI 90
Modelo saturado	30	,000	0			1,000				
Modelo 1	24	48,168	6	,000	8,028	,996	,982	,032	,024	,041
Modelo 2	24	106,408	6	,000	17,735	,992	,960	,050	,042	,059
Modelo 3	24	52,715	6	,000	8,786	,996	,980	,034	,026	,043
Modelo fijo	10	668,673	20	,000	33,434	,949	,924	,070	,065	,074

Para el análisis de la importancia de la primera aplicación como predictor en la materia de Matemáticas, los modelos estimados muestran ajustes muy buenos. En este caso, el modelo 3, es decir, el que mantiene fijas las covarianzas entre las cohortes 2 y 3 (los dos ciclos de Educación Secundaria) es el que obtiene mejores índices de ajuste con un valor de GFI igual a 0,999, AGFI 0,984 y RMSEA 0,023.

TABLA XVII. Importancia de la primera aplicación como predictor (Matemáticas)

Modelos	NPAR	CMIN	DF	P	CMIN/DF	GFI	AGFI	RMSEA	LO 90	HI 90
Modelo saturado	30	,000	0			1,000				
Modelo 1	27	21,126	3	,000	7,042	,998	,984	,030	,019	,043
Modelo 2	27	29,393	3	,000	9,798	,998	,979	,036	,025	,049
Modelo 3	27	13,994	3	,003	4,665	,999	,990	,023	,012	,036
Modelo fijo	24	44,354	6	,000	7,392	,997	,984	,031	,023	,040

<sup>(1)</sup> Recordemos que un valor no significativo de F sería el que nos llevaría a la conclusión de que las matrices de varianzas-covarianzas son invariantes.

Para finalizar, en el caso de Matemáticas, debemos tener precaución a la hora de afirmar la igualdad en la matriz de covarianzas ya que uno de los índices de ajuste analizados (RMSEA=0,07) sobrepasa ligeramente el límite para resultar aceptable.

## Discusión y conclusiones

Como hemos mencionado al comienzo de este artículo, existen varias causas que pueden explicar la variación en los patrones de correlaciones entre las variables de respuesta en estudios longitudinales como el que aborda esta investigación: cambios psicológicos, cambios organizativos del Sistema Educativo, aumento de la complejidad curricular e incluso artefactos asociados a la construcción de la escala común.

Una vez analizados los distintos modelos hipotetizados, parece evidente una mayor continuidad entre los datos provenientes de las cohortes 1ª y 2ª, es decir, entre Educación Primaria y primer ciclo de Educación Secundaria, que entre cualquiera de éstos y la cohorte 3ª.

Otro resultado relevante es la importancia como predictor de la primera aplicación, ya que es precisamente la correlación de esta variable con las de las aplicaciones sucesivas lo que se mantiene constante en los modelos que mejor ajustan.

Esta información nos lleva a pensar que la construcción de un modelo estadístico para describir y explicar el crecimiento de los centros escolares que permita seguir el progreso de unidades múltiples y anidadas puede ser el mismo para Primaria y primer ciclo de Secundaria. En cambio, para los dos ciclos de Educación Secundaria Obligatoria sería conveniente plantearse modelos diferenciados de crecimiento.

Asimismo y, enlazando con las características del contexto personal y académico de los alumnos objeto de estudio, podemos concluir que el cambio o transición de la Educación Primaria a la Educación Secundaria Obligatoria supone una variación significativa en las relaciones entre las variables dependientes medidas en ocasiones consecutivas, específicamente en el caso de la cohorte 3ª (segundo ciclo de Educación Secundaria), aunque este cambio no se manifiesta en los momentos inmediatos a la transición, sino que se hace visible cuando los alumnos encaran la parte final de la Educación Secundaria Obligatoria.

En consecuencia, cualquier estudio o evaluación que suponga una toma de datos longitudinal en los grupos de edad en los que se produce la transición entre la Educación Primaria y la Educación Secundaria debe tener muy en cuenta la posibilidad muy

sería de que la estructura correlacional entre las variables de respuesta experimente cambios de tal naturaleza que obliguen a plantearse modelos diferenciados para las cohortes en función de su proximidad al final del período obligatorio.

Esta variación asociada al tiempo se mantiene en los dos campos disciplinares que se han estudiado, Comprensión Lectora y Matemáticas, aunque cada uno de ellos tenga sus notas idiosincráticas.

Por último, y en cualquier caso, se mantiene la enorme importancia que las correlaciones de la primera aplicación con el resto tiene en cualquiera que sea el modelo y la cohorte analizados.

## Referencias bibliográficas

- ARBUCKLE, J. L. (2006). *Amos 7.0 user`s guide*. Chicago: SPSS, Inc.
- BAGOZZI, R. P. & YOUJAE, Y. (1988). On the evaluation of structural equation models. *Journal of the Academy of Marketing Science*, 16 (spring), 74-94.
- BALDWIN, C. K. & CADWELL, L. L. (2003). Development of the Free Time Motivation Scale for adolescents. *Journal of Leisure Research*, 35 (2), 129-151.
- BATISTA, J. M. Y COENDERS, G. (2000). *Modelos de educaciones estructurales*. Madrid: La Muralla.
- BLANCO, A., GONZÁLEZ, C. Y ORDÓÑEZ, X. (2009). Complejidad del rasgo evaluado y predicción del rendimiento escolar en diseños longitudinales: un estudio de simulación desde un enfoque multinivel. *Revista de Educación*, 348.
- BLOOM, B., HASTINGS, T. Y MADAUS, G. F. (1973). *Manual de evaluación formativa y acumulativa del aprendizaje del alumno*. Santiago de Chile: Centro de Perfeccionamiento, Experimentación e Investigaciones Pedagógicas (CPEIP).
- BOLLEN, K. A. & LONG, J. S. (eds.) (1993). *Testing structural equation models*. Newbury Park: Sage.
- (1949). A general distribution theory for a class of likelihood criteria. *Biometrika*, 36, 317-346.
- (1993). *Estadística para investigadores: introducción al diseño de experimentos, análisis de datos y construcción de modelos*. Barcelona: Reverté, D.L.
- BROWNE, M. W. Y CUDECK, R. (1993). Alternative ways of assessing model fit. En BOLLEN, K. A. Y LONG, J. S. (Eds.). *Testing structural equation models* (136-162). Newbury Park: Sage.

- BYRNE, B. M. (2001). *Structural equation modeling with AMOS: basic concepts, applications, and programming*. New Jersey: Lawrence Erlbaum Associates.
- CASTRO MORERA, M., RUIZ DE MIGUEL, C. Y LÓPEZ MARTÍN, E. (2009). Forma básica de crecimiento en los modelos de valor añadido: vías para la supresión del efecto de regresión y funciones de crecimiento no-lineales. *Revista de Educación*, 348.
- ELOSUA, P. (2005). Evaluación progresiva de la invarianza factorial entre las versiones original y adaptada de una escala de autoconcepto. *Psicothema*, 17 (2), 356-362.
- GAVIRÍA, J. L., BIENCINTO, CH., LÓPEZ, E., RUIZ DE MIGUEL, C., JORNET, J., CASTRO, M., Y GONZÁLEZ, C. (2007). Cuestiones metodológicas de la evaluación de sistemas educativos. *Symposium presentado al XII Congreso Nacional de Modelos de Investigación Educativa, San Sebastián*, 105-114.
- GAVIRIA, J. L. (2005). La equiparación del expediente de bachillerato en el proceso de selección de alumnos para el acceso a la Universidad. *Revista de Educación*, 337, 351-387.
- JÖRESKOG, K. G. & SÖRBOM, D. (1989). *LISREL 7, a guide to the program and applications*. Chicago: SPSS Publications.
- KIPLINGER, V. (2004). *Longitudinal study of student growth in reading, writing and mathematics achievement in thirtysix Colorado public schools districts*. Colorado: Academic school district twenty.
- HAIR, J., ANDERSON, R., TATHAM, R. Y BLACK, W. (1999). *Análisis Multivariante*. Madrid: Prentice Hall.
- INSTITUTO DE EVALUACIÓN (2006). *Sistema estatal de indicadores de la educación 2006*. Madrid: MEC
- INECSE (2003). *Evaluación de la educación secundaria obligatoria 2000*. Madrid: MECD.
- (2005). *Evaluación de la educación primaria 2003*. Madrid: MEC.
- LARA, F. (2004). *La escuela como compromiso*. Madrid: Editorial Popular.
- LIZASOAIN, L. Y JOARISTI, L. (2009). Análisis de la dimensionalidad en modelos de valor añadido: estudio de pruebas de matemáticas empleando técnicas factoriales y métodos no-paramétricos basados en TRI. *Revista de Educación*, 248.
- NAVARRO, E. Y REDONDO, S. (2007). Estudio sobre el Rendimiento en Matemáticas en España a partir de los Datos del Informe PISA 2003. Un Modelo Jerárquico de Dos Niveles. *Revista Electrónica Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, 5 (3), 118-136.
- PARDO, A. Y RUIZ, M.A. (2002). *SPSS11. Guía para el análisis de datos*. Madrid: McGraw Hill.

- SCHUMACKER, R. & LOMAX, R. (2004). *A beginner's to structural equation modelling*. Mahwah: Lawrence Erlbaum Associates.
- TABACHNICK, B. G. & FIDELL, L. S. (2001). *Using multivariate statistics*. Boston: Allyn and Bacon.
- TOPA, G., MORALES, J. F. & CAIERO, L. (2007). Social comparison and perceived breach of psychological contract: their effects on burnout in a multigroup analysis. *The Spanish Journal of Psychology*, 10 (1), 122-130.
- ZVOCH, K. & STEVENS, J.J. (2003). A multilevel, longitudinal analysis of middle school math and language achievement. *Education Policy Analysis Archives*, 11(20).

**Dirección de contacto:** Chantal Biencinto López. Universidad Complutense de Madrid. Facultad de Educación. Departamento de Métodos de Investigación y Diagnóstico en Educación. Avenida del Rector Royo Villanova s/n. 28040 Madrid. E-mail: alameda@edu.ucm.es



# **Análisis de la dimensionalidad en modelos de valor añadido: estudio de las pruebas de matemáticas empleando métodos no paramétricos basados en TRI**

## **Analysis of dimensionality at value-added models: a study of math test using non parametric methods based on item response theory**

Luis Lizasoain Hernández

Luis Joaristi Olariaga

*Universidad del País Vasco-Euskal Herriko Unibertsitatea. Departamento de Métodos de Investigación y Diagnóstico en Educación (MIDE). Donostia – San Sebastián, España.*

### **Resumen**

La evaluación basada en el valor añadido implica una metodología en la que se asumen supuestos fuertes. Se hace uso de diseños longitudinales y las puntuaciones de cada una de las mediciones deben ser situadas en una escala común. Ese escalamiento común supone un sólido diseño de equiparación. La única forma para asegurar la posibilidad de comparación dentro de los diseños longitudinales es emplear un modelo de medición donde toda escala comparte las mismas propiedades métricas. La posibilidad de la estimación en una escala común se basa en el supuesto de independencia local. El objetivo de este trabajo es analizar la estructura dimensional de las pruebas empleadas para evaluar el rendimiento académico en la asignatura de matemáticas en el contexto de la evaluación realizada en la Comunidad de Madrid en los cursos académicos 2005-06 y 2006-07 en tres cohortes: 5º y 6º de Educación Primaria, 1º-2º y 3º-4º de ESO. Para ello se evalúa su posible «unidimensionalidad» y la simplicidad o complejidad de su estructura empleando métodos no paramétricos basados en la Teoría de la Respuesta Ítem (TRI). Los resultados confirman estructuras esencialmente unidimensionales.

A su vez, se comprueba que el aumento de la complejidad de los contenidos implica un aumento en la complejidad de la estructura dimensional de las pruebas. Los modelos de medida utilizados son ciertamente bastante robustos frente a las violaciones leves de los supuestos relacionados con la independencia local.

*Palabras clave:* dimensionalidad, DIMTEST, DETECT, escalamiento vertical, valor añadido.

### **Abstract**

The added value based assessment implies the use of a methodology where strong assumptions are assumed. On one hand, longitudinal designs are used and, on the other hand, the scores of every time-moment have to be recoded into a common scale. This common scaling implies a robust equating design. The only way to ensure the comparability within the longitudinal designs is to use a measurement model where every scale shares the same metric properties. The possibility of common scale estimates is based upon the local independence assumption. The aim of this paper is to analyze the dimensional structure of a set of mathematics achievement tests used to assess the academic achievement in the Community of Madrid during 2005-06 and 2006-07 years and in three cohorts corresponding to the following academic degrees: 5th-6th of Primary Education, 1st-2nd and 3-4th of Obligatory Secondary Education. Their essential unidimensionality and the level of simplicity/complexity of their structure are assessed using nonparametric IRT-based procedures. These results confirm that the most of the tests have an essential unidimensional structure. In turn, the more of the contents complexity, the more of the tests dimensional structure complexity. The measurement models that have been used are certainly quite robust to minor violations related to the local independence assumptions.

*Key Words:* Dimensionality, DIMTEST, DETECT, vertical scaling, added value.

## **Introducción**

La evaluación basada en el valor añadido implica una metodología en la que se asumen algunos supuestos fuertes. Por un lado, se hace uso de diseños longitudinales en los que cada individuo es medido en varias ocasiones, y por otro, estas puntuaciones deben ser situadas en una escala común a todas ellas. Ese escalamiento común supone un sólido diseño de equiparación. Nos encontramos así con dos restricciones concurrentes. Las pruebas que se usan en cada ocasión deben tener validez curricular, es

decir, su contenido debe reflejar en cada momento los propios del curso que se evalúa; por otro, las puntuaciones obtenidas deben estar en la misma métrica común para todas las aplicaciones. La violación de estos supuestos puede tener consecuencias importantes para la validez de los resultados (Gaviria y Ruiz de Miguel, 2007).

Dado que se trata de una evaluación longitudinal, las pruebas tienen que adaptarse al currículo de cada año. Los diseños de valor añadido se utilizan fundamentalmente para la realización de comparaciones a lo largo del tiempo, entre e intra centros. El garante de la «comparabilidad» dentro de los diseños longitudinales es el modelo de medida empleado que asume la métrica común de las escalas de medida. La posibilidad de la estimación en una escala común se basa en el supuesto de independencia local.

La conjugación de la complejidad curricular y el supuesto de independencia local se ven comprometidos cuando la evaluación abarca varias cohortes de distintas edades. La utilización de un modelo común de valor añadido en cohortes de niveles educativos distintos, por ejemplo Primaria y Secundaria, supone asumir previamente que se tiene la misma estructura dimensional de las pruebas en cada una de las cohortes.

Según Stout et al. (1996) la evaluación de la dimensionalidad de una prueba se da en dos tipos de problema: por un lado, la verificación o refutación de la unidimensionalidad y por otro, la consiguiente descripción de la estructura multidimensional de la prueba, si ésta fuese necesaria. La verificación de la unidimensionalidad es imprescindible en aquellos procedimientos psicométricos que presuponen que los datos se ajustan a un modelo latente unidimensional o, cuando menos, que el desajuste respecto de esta condición no invalida la utilización de un procedimiento específico.

Las soluciones más apropiadas según Martínez Arias et al. (2006) empiezan por llevar a cabo un análisis factorial lineal paramétrico sobre la matriz de correlaciones entre ítems, habitualmente sobre la matriz de correlaciones tetracóricas, analizándose la dimensionalidad en función de distintos criterios: valores propios superiores a la unidad, diagrama de sedimentación, cociente entre los dos primeros valores propios, etc. Pero también esta cuestión se puede abordar también mediante métodos no paramétricos, en los que no se parte de ningún supuesto sobre las funciones de respuesta. Los procedimientos HCA/CCPROX, DETECT y DIMTEST (disponibles en el programa DIMPACK) están basados en el concepto de unidimensionalidad esencial de Stout y en la estimación de las covarianzas condicionales de cada par de ítems.

Desde esta perspectiva, el objetivo de este trabajo es analizar, mediante dichos métodos no paramétricos, la estructura dimensional de las pruebas empleadas para evaluar el rendimiento académico en la asignatura de matemáticas en el contexto de la

evaluación realizada en la Comunidad de Madrid en los cursos académicos 2005-06 y 2006-07 en tres cohortes: 5º y 6º de Educación Primaria, 1º- 2º y 3º- 4º de ESO. Los datos utilizados para la realización de este trabajo proceden del Proyecto de I+D con referencia SEC2003-09742, ya finalizado, titulado: *El valor añadido en educación y la función de producción educativa: un estudio longitudinal*. Se ha evaluado la posible unidimensionalidad y la simplicidad o complejidad de la estructura de las pruebas.

Como se puede ver en la Tabla I, para los alumnos que en 2005-06 entraron en 5º de Primaria (cohorte 1ª) se han recogido datos en octubre de 2005 (primera medición), y junio de 2006 (segunda medición). En el curso 2006-07, cuando estos alumnos estaban en 6º de Primaria, se recogió información en noviembre de 2006 (tercera medición) y junio de 2007 (cuarta medición). Con las cohortes 2ª (correspondiente al primer ciclo de Secundaria) y 3ª (correspondiente al segundo ciclo de Secundaria) se ha procedido de manera análoga.

TABLA I. Tamaño de la muestra por cohortes y aplicaciones

Cohorte	Curso	Octubre 2005	Junio 2006	Curso	Noviembre 2006	Junio 2007
1ª	5º Primaria	4211	4173	6º Primaria	3627	3757
2ª	1º ESO	5106	4882	2º ESO	3327	3403
3ª	3º ESO	4736	4272	4º ESO	2761	2772

Las variables dependientes utilizadas han sido las puntuaciones obtenidas en las pruebas de Matemáticas elaboradas *ad hoc* aplicadas a los alumnos de los centros de la muestra.

## Análisis mediante métodos no paramétricos basados en la TRI

### Introducción

Como al inicio se ha mencionado, además de los enfoques basados en las técnicas factoriales, la evaluación de la dimensionalidad de las pruebas es posible abordarla también con procedimientos no paramétricos basados en la Teoría de Respuesta al Ítem. En concreto aquí vamos a presentar la aplicación a nuestras pruebas de los procedimientos DETECT y DIMTEST que forman parte del paquete DIMPACK elaborado por

*The Roussos-Stout Software Development Group y distribuido por Assessment Systems Corp (2006).*

Estos procedimientos se basan en el concepto de «unidimensionalidad esencial» (Stout, 1987, 1990) y en la covarianza condicional de pares de ítems (Stout et al., 1996). Este principio de dimensionalidad esencial se basa a su vez en la distinción establecida por McDonald (1981) entre los principios «fuerte» y «débil» de independencia local (Strong Local Independence, SLI; Weak Local Independence, WLI).

Stout define la dimensionalidad esencial de una prueba como el número mínimo de dimensiones del espacio de rasgos latentes tal que produzca un modelo que sea localmente independiente y monótono. Conceptualmente, la idea es que una prueba es «esencialmente unidimensional» si la puntuación de la misma está en función de «un» rasgo latente dominante y también de uno o varios rasgos menores, secundarios o de ruido (*nuisance latent traits*). Una descripción más formalizada se encuentra en Zhang y Stout (1999), y en Jang y Roussos (2007).

## Procedimiento DETECT

DETECT es un procedimiento cuya finalidad es estudiar la posible multidimensionalidad de una prueba. Para ello obtiene el índice DETECT que expresa la «cantidad» de multidimensionalidad de la misma. Esto, en términos geométricos, hace referencia a la dispersión de los vectores de los ítems. En cualquier caso, no debe ser confundido con el número de dimensiones de una prueba. Este procedimiento está basado en que, dada una prueba multidimensional, los ítems que miden la misma dimensión tendrán covarianzas condicionales positivas, mientras que para los que midan otras dimensiones distintas, sus covarianzas serán negativas de forma que el citado índice puede ser considerado como una media ponderada (Stout, W., Nandakumar, R. y Habing, R., 1996, Roussos y Ozbek, 2006 y Stout et al., 2006). Para una descripción rigurosa de la fundamentación matemática de este índice puede recurrirse al trabajo de Zhang (2007).

El procedimiento DETECT puede emplearse tanto desde un planteamiento exploratorio como confirmatorio. En el primer modo, DETECT busca cuál es la partición que maximiza el índice y asigna ítems a los conglomerados. En modo confirmatorio una submuestra se emplea para la parte exploratoria y otra para validar la partición, es decir, para validar el modelo propuesto por el usuario (Jang y Roussos, 2007). En general, si se dispone de muestras de tamaño suficiente, se recomienda usar el procedimiento de validación cruzada (Monahan et al., 2007).

Siguiendo estas pautas, hemos aplicado el procedimiento DETECT a las 24 pruebas que estamos estudiando después de haber dividido la muestra total en dos submuestras de similar número de efectivos con objeto de que una opere como submuestra de exploración y la otra de validación. En la Tabla II, que presentamos a continuación, aparecen los resultados de este proceso.

TABLA II. Aplicación de DETECT con validación cruzada a las 24 pruebas

Cohorte	Octubre 2005		Junio 2006		Octubre 2006		Junio 2007		Medias
	A	B	A	B	A	B	A	B	
5°-6°	DETECT:	DETECT:	DETECT:	DETECT:	DETECT:	DETECT:	DETECT:	DETECT:	
	0.0892	0.0966	0.1261	0.1332	0.1997	0.2398	0.1113	0.1575	0,1442
	IDN :	IDN :	IDN :	IDN :	IDN :	IDN :	IDN :	IDN :	
	0.5513	0.5385	0.5692	0.5756	0.5775	0.5978	0.5377	0.5718	0,5649
	Ratio:	Ratio:	Ratio:	Ratio:	Ratio:	Ratio:	Ratio:	Ratio:	
	0.2234	0.2215	0.2270	0.2663	0.3058	0.3577	0.1918	0.2306	0,2530
Max: 4	Max: 4	Max: 4	Max: 3	Max: 4	Max: 4	Max: 4	Max: 4	3,8750	
1°-2°	DETECT:	DETECT:	DETECT:	DETECT:	DETECT:	DETECT:	DETECT:	DETECT:	
	0.1137	0.2034	0.1187	0.1762	0.1686	0.0926	0.1567	0.0993	0,1412
	IDN :	IDN :	IDN :	IDN :	IDN :	IDN :	IDN :	IDN :	
	0.5601	0.6248	0.5548	0.5968	0.5965	0.5439	0.5526	0.5377	0,5734
	Ratio:	Ratio:	Ratio:	Ratio:	Ratio:	Ratio:	Ratio:	Ratio:	
	0.2084	0.3468	0.1895	0.3007	0.2610	0.1367	0.2312	0.1472	0,2277
Max: 3	Max: 5	Max: 4	Max: 4	Max: 5	Max: 4	Max: 5	Max: 4	4,2500	
3°-4°	DETECT:	DETECT:	DETECT:	DETECT:	DETECT:	DETECT:	DETECT:	DETECT:	
	0.2966	0.1205	0.2292	0.1568	0.2337	0.2426	0.1683	0.1662	0,2017
	IDN :	IDN :	IDN :	IDN :	IDN :	IDN :	IDN :	IDN :	
	0.6628	0.5692	0.6323	0.5833	0.5868	0.5946	0.5505	0.5420	0,5902
	Ratio:	Ratio:	Ratio:	Ratio:	Ratio:	Ratio:	Ratio:	Ratio:	
	0.4968	0.2009	0.4947	0.2529	0.2979	0.2739	0.1947	0.2041	0,3020
Max: 4	Max: 4	Max: 5	Max: 4	Max: 4	Max: 4	Max: 4	Max: 5	4,2500	

El contenido de cada casilla es como sigue:

- DETECT: es el índice del mismo nombre y ha de interpretarse como un indicador de la cantidad de multidimensionalidad. Inmediatamente vamos a exponer los criterios de interpretación de este índice.

- IDN: es un índice que informa del grado de simplicidad o complejidad de la estructura, de forma que valores cercanos a 1 denotan una estructura simple.
- Ratio: es la proporción entre el valor del índice DETECT obtenido y su valor máximo posible. Al igual que en el caso anterior, cuanto más cercano a 1, más simple se puede considerar la estructura.
- Max: indica el número de particiones de los ítems que hacen máximo el valor del índice DETECT.

Veamos con más detalle cada uno de estos parámetros. Como en su momento se apuntó, el procedimiento comienza buscando el número de particiones que hacen máximo el índice DETECT. El valor apuntado, valorado conjuntamente con los demás, es un buen indicador del tipo de estructura con que nos encontramos. Así vemos que en nuestro caso, en la gran mayoría de las pruebas (17) el número de particiones que maximizan el índice es de cuatro, en cinco casos son cinco; y sólo en dos el máximo está en tres. El índice DETECT es una media ponderada de todas las covarianzas condicionales de forma que el peso será +1 cuando los dos ítems provengan de un mismo conglomerado, y -1 si la pareja procede de distintos conglomerados (Jang y Roussos, 2007). En relación a los criterios de interpretación del mismo, en la literatura nos encontramos con varios, aunque básicamente se pueden resumir en los dos que presentamos en la Tabla III (Monahan et al., 2007).

TABLA III. Tamaño de la muestra por cohortes y aplicaciones

Unidimensionalidad esencial	0,0 - 0,1		
Multidimensionalidad débil	0,1 - 0,5	Unidimensionalidad esencial	0,0 - 0,2
Multidimensionalidad moderada	0,5 - 1,0	Multidimensionalidad de débil a moderada	0,2 - 0,4
Multidimensionalidad fuerte	1,0 - 1,5	Multidimensionalidad de moderada a grande	0,4 - 1,0
Multidimensionalidad muy fuerte	Sup. 1,5	Multidimensionalidad muy fuerte	Sup. 1,0

En nuestro caso, el optar por uno u otro no es baladí porque, como puede verse, la mayoría de las pruebas obtienen valores que se encuentran en los límites de ambos criterios (17 pruebas tienen un índice con un valor comprendido entre 0,11 y 0,18; y el promedio de las 24 es 0,1624) lo que supone que, si se adopta el primer criterio, la mayoría de las pruebas podrían considerarse como esencialmente unidimensionales; mientras que si optamos por el segundo, hablaríamos más bien de una débil multidimensionalidad.

Con respecto al segundo índice, el valor del IDN se basa en la proporción de signos positivos y negativos de las covarianzas condicionales con respecto al máximo

posible que se da cuando todos los signos «intra-conglomerados» son positivos y todos los «entre-conglomerados» son negativos. Este índice ha de interpretarse, por tanto, como un indicador de hasta qué punto la estructura puede considerarse como simple o compleja. Se entiende como estructura simple perfecta aquella en que los ítems se agrupan en  $k$  dimensiones que son subconjuntos disjuntos. Obviamente no existen estructuras simples perfectas (que darían un valor de uno en este índice), sino que esta característica se puede considerar como un continuo que va desde cero (máxima complejidad) hasta uno (máxima simplicidad).

Y con respecto al tercer parámetro, ya hemos anticipado que es la ratio entre el valor del índice DETECT efectivamente obtenido y el máximo que puede alcanzar. Su interpretación es similar al anterior.

A modo de resumen final, del examen de los promedios por cohortes podemos concluir que -en conjunto- nos encontramos con unas pruebas de multidimensionalidad débil o incluso en algunos casos esencialmente unidimensionales. Si aplicamos el criterio más riguroso nos encontramos con 4 de ellas unidimensionales (índice menor de 0,1). Junto a éstas hay 14 cuyo valor está comprendido entre 0,10 y 0,20 lo que supone, dependiendo del criterio que se emplee, que se calificarían bien como unidimensionales, bien de débil multidimensionalidad.

Lo que sí hay que señalar es que el comportamiento de este índice no es igual en las tres cohortes pues la de 3º-4º tiene una media claramente superior. En lo que no hay tanta variación (aunque la tercera cohorte también arroja el valor más alto) es en lo relativo al índice de complejidad cuyo promedio global es 0,576 lo que nos indica que -al margen de que las pruebas sean o no esencialmente unidimensionales- la estructura es de una cierta complejidad. Este resultado es también claramente coincidente con la estructura factorial que hemos encontrado en la que, como vimos, para todas las pruebas, prácticamente la gran mayoría de los ítems carga en el primer factor; mientras que, con respecto al segundo, son pocos los ítems que saturan en el mismo sin hacerlo simultáneamente en el primero.

Esto parece indicar que nos encontramos ante una estructura dimensional compuesta muy probablemente por una dimensión mayor, principal, y una o varias secundarias o menores con incluso algunas compuestas por sólo un ítem. Es una estructura típica que Stout caracterizó al formular la unidimensionalidad esencial señalando además -como ocurre en nuestro caso- que «una estructura simple no implica que las dimensiones no estén correlacionadas» (Stout, 1996, p. 332). Como ahora veremos, el análisis de conglomerados jerárquico apunta también en este sentido.

## Procedimiento DIMTEST

El procedimiento DIMTEST (Stout et al., 1996; Nandakumar y Stout, 1993; Stout, 1987) se basa en comparar la covarianza condicional de dos subconjuntos de ítems de una prueba (el subconjunto de evaluación, AT; y el de partición, PT). Para ello contrasta la hipótesis nula de que AT es de similar dimensionalidad que PT, frente a la hipótesis alternativa de que AT es dimensionalmente homogéneo y distinto que PT.

El principio de esta comparación radica en que las covarianzas condicionales de DIMTEST son las covarianzas entre los ítems de AT condicionadas al rasgo latente  $\xi$  siendo medido por PT. La idea es que si los ítems de AT y los de PT miden la misma dimensión, entonces las covarianzas condicionales entre los ítems de AT serán igual a cero, mientras que si AT mide una dimensión distinta que PT, entonces las covarianzas condicionales serán positivas (Jang y Roussos, 2007). El contraste se lleva a cabo mediante el estadístico T, y los detalles relativos al mismo tanto en la versión inicial que requería dos subconjuntos AT como en la actual que sólo emplea uno, se pueden encontrar en Stout, Froelich y Gao (2001, p. 5-14) y en Stout, Nandakumar y Habing (1996, p. 44-47).

Este principio así formulado hace que la elección de los ítems que conforman AT sea crucial y que determine en gran medida el resultado de la prueba estadística. Esto ya fue observado por Wang y Hocevar (1994) en un trabajo en el que precisamente estudiaban la dimensionalidad de pruebas de matemáticas cuando afirmaron que el grado de dimensionalidad esencial depende de la elección de ítems de AT. A este respecto, los criterios de selección de los ítems que conforman AT son formulados de la siguiente manera (Stout et al., 1996; Stout, Froelich y Gao, 2001):

- Los ítems de AT han de ser relativamente homogéneos.
- Los ítems de AT han de ser lo más distintos con respecto a los de PT.
- El número de ítems de PT ha de ser lo suficientemente grande (al menos 15), y el de AT no más de un tercio de los de PT con un mínimo de tres.

Es decir, dado el contraste que DIMTEST realiza, se trata de seleccionar como miembros pertenecientes al subconjunto AT aquellos ítems que conformen un aglomerado lo más unidimensional posible y en una dimensión distinta del resto de los ítems de la prueba (que conforman PT). De esta manera si, a priori, la selección es exigente y AT es lo más unidimensional posible y lo más distinto posible de PT, podemos estar razonablemente seguros de que si el valor de T tiene una probabilidad mayor

que el nivel alfa, si se acepta la hipótesis nula, entonces la prueba en su conjunto se puede considerar como esencialmente unidimensional dado que a pesar de haber maximizado a priori la diferencia entre ambos subconjuntos, la misma no ha resultado finalmente significativa.

Y los métodos propuestos para encontrar ítems que satisfagan estos criterios son el análisis factorial exploratorio, el análisis de conglomerados jerárquico y el juicio de expertos basado en el contenido, el dominio cognitivo y la estructura de la prueba. Y en esta fase del trabajo, nosotros hemos aplicado estos tres criterios. Veamos los resultados de cada uno de ellos.

Básicamente DIMTEST ofrece dos procedimientos para seleccionar los ítems de AT: o son definidos por el usuario en función de algún criterio de los anteriormente señalados, o bien se emplea un algoritmo incorporado que realiza esta tarea por nosotros. En primer lugar, probamos a emplear como criterio los conglomerados generados por DETECT o por HCC. En el primer caso, como ya comentamos, existe la posibilidad de especificar como parámetro del procedimiento el número de dimensiones, de conglomerados a extraer. Pues bien, si este valor lo fijamos en dos, obtenemos una partición de los ítems en dos subconjuntos disjuntos que podemos emplear (en función de su tamaño relativo) como AT y PT. De manera similar, si para cada prueba realizamos un análisis de conglomerados jerárquico mediante el procedimiento HCC y seleccionamos el penúltimo paso (la solución con dos conglomerados) podemos operar de la misma forma.

Pues bien, en ambos casos y para la mayoría de las 24 pruebas, los resultados de DIMTEST son los que cabía esperar: para aquellas pruebas en que DETECT había arrojado un valor del índice que denotaba unidimensionalidad esencial, el contraste de T mediante DIMTEST resultaba no significativo, mientras que en las que el índice era alto, la prueba resultaba estadísticamente significativa. Por último, para los valores intermedios del índice DETECT (de dimensionalidad débil), los resultados variaban.

Desde este punto de vista, podemos considerar que, en nuestro caso, el empleo de DIMTEST después de DETECT o HCC resultaba redundante y no nos aportaba información adicional relevante.

A continuación, optamos por que fuera el propio programa el que realizase la selección de los ítems de AT. El criterio de selección que el programa incorpora emplea los resultados de un análisis exploratorio de componentes principales, de forma que, de la solución no rotada, selecciona aquellos ítems cuyas saturaciones son altas en el segundo factor y bajas en el primero. Esto en principio es un procedimiento muy coherente con la estrategia analítica de DIMTEST. Véase Tabla IV.

**TABLA IV. Resultados DIMTEST, método automático AFE**

P (AFE)	Octubre 2005		Junio 2006		Octubre 2006		Junio 2007	
	A	B	A	B	A	B	A	B
5°-6°	0.0032	0.0125	0.0001	0.0000	0.0000	0.0000	0.1387	0.0190
1°-2°	0.0046	0.0010	0.0000	0.0000	0.4708	0.0000	0.4434	0.1412
3°-4°	0.0000	0.0015	0.0485	0.0000	0.0000	0.0036	0.0075	0.0002

Como vemos, los resultados aquí podrían, en principio, considerarse muy divergentes con respecto a los obtenidos mediante DETECT. Aquí sólo seis de las 24 pruebas resultan esencialmente unidimensionales, y eso empleando como nivel alfa el 0,05. Si optamos por el 0,01 se nos quedan en cuatro. Por el contrario, recordemos cómo eran 18 las pruebas que tenían un índice DETECT inferior a 0,2. Así vistos los resultados, la divergencia entre ambos procedimientos parece grande, pero la cuestión toma otro cariz si en el caso de DETECT optamos por un criterio más exigente, pues entonces sólo son cuatro las pruebas con índice inferior a 0,10. Desde esta perspectiva parecería que los enfoques convergen, aunque si examinamos las tablas con detalle veremos que es sólo a nivel numérico global pues son pocas las pruebas cuyos resultados coinciden bajo ambos criterios (más adelante resumiremos esto en una tabla global).

En nuestra opinión, esta divergencia entre ambos enfoques puede explicarse si nos detenemos a examinar el modo en que este criterio «automático» de DIMTEST procede. Como hemos dicho, se seleccionan los ítems del segundo factor no rotado de ACPY lo que ocurre es que, como vimos antes, la gran mayoría de estas pruebas tienen una estructura compleja y esto hace que el criterio adoptado por DIMTEST no funcione bien pues no hay casi ítems que claramente carguen en el segundo factor sin hacerlo simultáneamente en el primero. Este hecho es señalado por Finch (2006) al indicar que habitualmente las soluciones iniciales no reflejan una estructura simple. Con objeto de afrontar este problema, hemos optado por un criterio que emplee las soluciones factoriales, pero no del ACP sin rotar, sino los resultados del análisis factorial de información completa con solución rotada según el criterio PROMAX que anteriormente hemos expuesto.

Pensamos que este enfoque permite satisfacer mejor los criterios que DIMTEST recomienda a la hora de seleccionar ítems para AT pues la solución rotada permite seleccionar ítems que sólo carguen en el segundo factor sin hacerlo el primero. Es cierto que al tratarse de una solución oblicua, ambos factores están relacionados, pero es que la estructura de las pruebas parece indicar la existencia de una dimensión principal que estaría representada por el primer factor y una dimensión secundaria de mayor o menor importancia relativa en las diferentes pruebas que sería el segundo factor. (Hemos probado también soluciones con rotación ortogonal como VARIMAX

pero ocurre prácticamente lo mismo que con las soluciones iniciales no rotadas). En concreto hemos procedido de la siguiente manera.

En primer lugar procedimos para cada prueba a la selección de los ítems de AT empleando las cargas factoriales de la solución PROMAX obtenida mediante TES-FACT. En aplicación de los criterios de DIMTEST, se seleccionaron para cada prueba aquellos ítems cuyas cargas factoriales fuesen iguales o superiores a 0,35 en el segundo factor y menores que 0,15 en el primero (Finch, 2006).

Las Tablas V y VI muestran respectivamente los valores de la probabilidad, y el número, contenido y dominio cognitivo de los ítems seleccionados como miembros de AT en aplicación del criterio anterior.

TABLA V. Resultados DIMTEST, AT manual mediante saturaciones PROMAX

P (Promax)	Octubre 2005		Junio 2006		Octubre 2006		Junio 2007	
	A	B	A	B	A	B	A	B
5° - 6°	0.4006	0.0789	0.0008	0.0001	0.1236	0.0000	0.3901	0.2039
1° - 2°	0.2849	0.1074	0.0180	0.0212	0.0002	0.0145	0.0905	0.2493
3° - 4°	0.0002	0.0000	0.0782	0.0234	0.0191	0.0019	0.1562	0.0007

TABLA VI. Descripción ítems solución anterior

Ítems	Octubre 2005		Junio 2006		Octubre 2006		Junio 2007		
	A	B	A	B	A	B	A	B	
5°-6°	Con.	Nu:4	Nu:4	Nu:0	Nu:0	Nu:6	Nu:0	Nu:3	Nu:1
		Ge:0	Ge:0	Ge:4	Ge:5	Ge:1	Ge:7	Ge:1	Ge:2
		Es:0	Es:0	Es:0	Es:0	Es:1	Es:2	Es:2	Es:2
		Me:0	Me:0	Me:0	Me:1	Me:3	Me:0	Me:1	Me:2
	Dom.	Con:1	Con:2	Con:4	Con:5	Con:5	Con:4	Con:5	Con:3
		Apr:3	Apr:2	Apr:0	Apr:1	Apr:6	Apr:5	Apr:2	Apr:4
	N	4	4	4	6	11	9	7	7
1°-2°	Con.	Nu:3	Nu:1	Nu:4	Nu:8	Nu:7	Nu:1	Nu:5	Nu:2
		Ge:1	Ge:2	Ge:3	Ge:3	Ge:0	Ge:4	Ge:1	Ge:1
		Me:0	Me:0	Me:0	Me:0	Me:0	Me:1	Me:0	Me:0
		Ta:1	Ta:0	Ta:0	Ta:1	Ta:0	Ta:0	Ta:0	Ta:1
	Dom.	Con:4	Con:1	Con:2	Con:5	Con:6	Con:2	Con:1	Con:2
		Apr:1	Apr:2	Apr:5	Apr:7	Apr:1	Apr:4	Apr:5	Apr:2
	N	5	3	7	12	7	6	6	4
3°-4°	Con.	Nu:2	Nu:4	Nu:0	Nu:2	Nu:7	Nu:2	Nu:1	Nu:3
		Ge:0	Ge:0	Ge:0	Ge:0	Ge:0	Ge:1	Ge:4	Ge:0
		Fu:1	Fu:0	Fu:3	Fu:0	Fu:0	Fu:2	Fu:1	Fu:3
		Es:0	Es:1	Es:0	Es:1	Es:0	Es:2	Es:0	Es:0
	Dom.	Con:1	Con:2	Con:0	Con:1	Con:2	Con:0	Con:0	Con:3
		Apr:2	Apr:3	Apr:3	Apr:2	Apr:5	Apr:7	Apr:6	Apr:3
	N	3	5	3	3	7	7	6	6

La solución aportada por DIMTEST para estos AT es la que finalmente adoptamos y del examen de la misma vemos que nos encontramos con once pruebas que podemos considerar como básicamente unidimensionales. Tal y como antes comentamos, al final de este trabajo compararemos todos los criterios y procedimientos que hemos empleado con objeto de triangular y proponer una solución final global, por lo que ahora nos vamos a limitar a señalar cómo las dos primeras cohortes tienen una proporción mayor de pruebas unidimensionales, mientras que la cohorte de 3º y 4º de la ESO se puede considerar como no unidimensional lo que de nuevo viene a confirmar lo comentado anteriormente.

Pero ahora nos interesa centrarnos en los ítems seleccionados como miembros de AT. Por término medio, la aplicación de los criterios señalados supone que el tamaño medio de este subconjunto es de 5,8 ítems (5,45 para las pruebas unidimensionales y 6,3 en el caso de las otras). En consecuencia, la proporción AT-PT es de aproximadamente uno a seis o siete dado que el número de ítems de las pruebas oscila entre 36 y 40.

Estos valores pensamos que son un buen indicador del tipo de estructura que nos encontramos, pero además de esto nos interesa también fijarnos en el contenido sustantivo de los mismos. Según la matriz de especificaciones elaborada por los diseñadores de las pruebas (inspectores), los contenidos y dominios cognitivos son los que aparecen en la Tabla VI. Los contenidos son: números, geometría, estadística, medida, tablas y funciones; y los dominios cognitivos son conocimientos (Con) y aplicación o resolución de problemas (Apr). En todas las pruebas los ítems mayoritarios pertenecen al contenido «numérico-aritmético» y al dominio cognitivo de conocimientos.

Y del examen de la Tabla VI podemos ver que en gran parte de los casos nos encontramos que este subconjunto AT está formado mayoritariamente bien por ítems de contenidos distintos al numérico (geometría, estadística, medida, etc.) y/o por aquellos ítems que suponen aplicación o resolución de problemas. Esta tendencia no es unánime ni homogénea pero sí mayoritaria tanto para las pruebas unidimensionales como para las restantes. Vamos a ver con algo de detalle un caso paradigmático como es el de la forma B de la prueba aplicada a 6º de Primaria en octubre de 2006, pero no hay que olvidar que junto a un caso tan claro como éste hay notables excepciones que al final enumeraremos.

El valor de T asociado a esta prueba es de 6,64 lo que supone una probabilidad menor que 0,0001. Es decir, nos encontramos con una prueba claramente no unidimensional desde la perspectiva de DIMTEST. Los nueve ítems que conforman el subconjunto AT (1, 3, 5, 6, 7, 8, 9, 10 y 11) son dos de estadística y nueve de geometría, mientras que desde el punto de vista del dominio se reparte a medias entre conocimientos

y resolución de problemas. Los 30 ítems restantes (que componen PT) son mayoritariamente de contenido numérico (16) o relacionados con la medida y las magnitudes (diez) habiendo también tres de geometría y una de estadística.

Para mayor evidencia en este sentido, nos encontramos además que uno de los conglomerados propuestos por el procedimiento DETECT está compuesto básicamente por estos ítems y lo mismo ocurre con la solución de cinco conglomerados de HCC en que éstos conforman el quinto grupo. (A este respecto, pensamos que de nuevo es síntoma de complejidad estructural el hecho de que sea en el quinto grupo cuando surja esta estructura y no en la de dos conglomerados lo que nos ratifica en lo que antes apuntamos: dada esta complejidad las soluciones de dos conglomerados proporcionadas por DETECT o HCC no son satisfactorias).

En conclusión, parece claro en este caso que nos encontramos con una prueba no unidimensional compuesta por una primera dimensión principal en la que saturan la mayoría de los ítems y una secundaria centrada fundamentalmente en ítems de geometría.

Como antes apuntábamos, esta caracterización de los dos factores puede encontrarse en muchas de las pruebas (sean o no esencialmente unidimensionales), pero hay también excepciones. Y entre ellas destacamos la forma B de 5º de primaria aplicada en octubre de 2005, la forma A de 6º de primaria de octubre de 2006 (la que debería ser equivalente a la que acabamos de describir y que al menos dimensionalmente no lo es), la forma A de 6º de primaria aplicada en junio de 2007, la forma A de 1º de la ESO de octubre de 2005 y, por último, la forma A de 2º de la ESO de octubre de 2006. Queremos resaltar el caso de ésta última porque, además de no ajustarse al patrón descrito, es además no unidimensional. Se trata de cinco claras excepciones sobre un total de 24, por lo que la existencia de un único patrón debe formularse con las debidas precauciones.

Stout, Nandakumar y Habing (1996); Perkhounkova y Dunbar (1999); y Walker et al. (2006), al analizar pruebas de matemáticas llevan a cabo planteamientos similares combinando el análisis de contenido de ítems y pruebas con los resultados de DIMTEST. Hattie et al. (1996) concluyen su trabajo afirmando que DIMTEST está basado en el principio débil de independencia local y está diseñado, no tanto para identificar si un conjunto de ítems es o no unidimensional, sino si hay una dimensión suficientemente dominante. Si además combinamos las perspectivas que nos proporcionan DETECT y DIMTEST pensamos que es posible lograr evidencias relativas a la posible unidimensionalidad esencial de las pruebas en el sentido que aquí hemos concluido.

## Comparación de resultados y conclusiones finales

A lo largo de este trabajo hemos abordado la dimensionalidad de las 24 pruebas desde diferentes perspectivas y con diferentes procedimientos y herramientas analíticas. En los distintos apartados hemos ido apuntando las conclusiones parciales y hemos ido igualmente señalando las concordancias entre criterios así como también las divergencias observadas. Para finalizar, hora es que recopilemos sintetizando resultados y apuntando las conclusiones finales.

Con objeto de evaluar la dimensionalidad de las pruebas, seis son los principales criterios que hemos empleado:

- Que en la solución factorial, sólo tenga un autovalor superior a uno el primer factor.
- Que el cociente entre los autovalores de los dos primeros factores sea superior a cinco.
- La significatividad de la T según el procedimiento DIMTEST habiendo seleccionado los ítems de AT según las cargas factoriales de la solución PROMAX.
- La significatividad de la T según el procedimiento DIMTEST habiendo seleccionado los ítems de AT el propio programa mediante AFE.
- Que el valor del índice DETECT sea inferior a 0,10.
- Que el valor del índice DETECT sea inferior a 0,15.

Pues bien, en la Tabla VII, siguiendo la presentación habitual de cohortes, aplicaciones y formas de las pruebas, se consigna en cada casilla los criterios que cada prueba satisfice.

TABLA VII. Síntesis de criterios

Examen Conjunto	Octubre 2005		Junio 2006		Noviembre 2006		Junio 2007	
	A	B	A	B	A	B	A	B
5° - 6°	3,5,6	3,4,5,6	2,6	2,6	2,3	2	1,2,3,4,6	1,2,3,4
1° - 2°	1,2,3,6	3	1,2,6	1,2	4	5,6	2,3,4	3,4,5,6
3° - 4°		1,6	3,4		1	1	1,3	1

A la vista de esto hemos clasificado las 24 pruebas en tres categorías atendiendo a su unidimensionalidad esencial y nos encontramos con:

- Diez pruebas de las que parece haber evidencia suficiente de unidimensionalidad esencial (casillas sombreadas en gris oscuro).

- Seis pruebas de multidimensionalidad débil o en las que la evidencia de unidimensionalidad esencial no es tan concluyente (casillas sombreadas en gris claro).
- Ocho pruebas de multidimensionalidad baja o moderada (casillas sin sombreado).

Los criterios para asignar a cada una de estas tres categorías son:

- Se consideran pertenecientes al primer grupo (esencialmente unidimensionales) si:
  - el valor del índice DETECT tiene un valor inferior a 0,10.
  - si DETECT es inferior a 0,15 y además satisfacen otros dos criterios más cualesquiera.
  - si el valor de T no ha resultado significativo en los 2 procedimientos DIMTEST.
- Se consideran de multidimensionalidad débil aquellos que satisfacen uno de los criterios tres o seis.
- Y el resto son considerados de multidimensionalidad baja o moderada.

Como vemos, la prelación de los criterios es clara. Si el índice DETECT es inferior a 0,10 la prueba se considera esencialmente unidimensional. La misma categoría tienen aquellas pruebas cuyo índice es menor que 0,15 y que además hayan satisfecho otros dos criterios. Y lo mismo ocurre si en los dos procedimientos DIMTEST, la T de Stout ha resultado no significativa. En definitiva, el criterio básico es el índice DETECT en su versión más estricta y en los demás casos se considera esencialmente unidimensional si se han obtenido evidencias adicionales satisfaciendo dos (si son DIMTEST) ó más de dos criterios. En este primer grupo se encuentran diez pruebas que como vemos pertenecen cuatro a la primera cohorte, cinco a la segunda y sólo una a la tercera. En el grupo intermedio se encuentran las seis pruebas que sólo han satisfecho dos criterios siendo éstos el dos o el tres. Y el tercer grupo es aquel en el que hay pocas o ninguna evidencia de unidimensionalidad. Se trata de ocho pruebas de las cuáles cinco pertenecen a la tercera cohorte. En general, son pruebas de baja o moderada multidimensionalidad pues no olvidemos que el índice DETECT es muy bajo en general. De las ocho pruebas no todas siguen un patrón idéntico:

- Podemos considerar como las de una multidimensionalidad relativa más clara las pruebas correspondientes a las aplicaciones de octubre 2005 (forma A), las dos formas de octubre de 2006 y la forma B de junio de 2007, todas ellas de la tercera cohorte.

- La aplicación de la forma A de octubre de 2006 de 2º de la ESO es peculiar pues resulta unidimensional en el procedimiento DIMTEST con AFE (criterio cuatro) con una probabilidad muy elevada, y en cambio sólo resulta unidimensional en dicho criterio.
- Todas las pruebas marcan en algún criterio, excepto la forma A de octubre 2005 y la B de junio de 2006 de la tercera cohorte.

Si examinamos la Tabla VIII por filas (cohortes), vemos que la primera (5º y 6º de primaria) resulta ser la más unidimensional, seguida de la de 1º y 2º de la ESO; y es, como ya veníamos comentando, la de 3º y 4º de la ESO la de multidimensionalidad más acusada.

**TABLA VIII.** Categorías de dimensionalidad

	Esencialmente unidimensional	Multidimensionalidad. débil	Baja o moderada multidimensionalidad
5º-6º	4	3	1
1º-2º	5	1	-
3º-4º	1	2	5

De los seis criterios estudiados, el menos fiable es el primero pues son nueve las pruebas que marcan en él y de las mismas, cuatro resultan de la tercera categoría, dos de la segunda y tres de la primera. En definitiva, no parece discriminar adecuadamente.

Por último, la Tabla IX muestra la simplicidad o complejidad relativa de las pruebas en función del índice IDN proporcionado por DETECT. Se han considerado de estructura simple (y sombreado en gris) aquellas pruebas con un índice superior a 0,58. Dentro de las mismas se ha consignado si, de acuerdo con la síntesis anterior, resultaban multidimensionales. Vemos como de las nueve que resultan de estructura más relativamente simple, siete pertenecen a la tercera categoría.

**TABLA IX.** Simplicidad-complejidad relativa

Estructura simple vs. Compleja	Octubre 2005		Junio 2006		Noviembre 2006		Junio 2007	
	A	B	A	B	A	B	A	B
5º - 6º						Multidim.		
1º - 2º				Multidim.	Multidim.			
3º - 4º	Multidim.			Multidim.	Multidim.	Multidim.		

## Conclusiones finales

Si a todo esto unimos a la caracterización de los dos factores que hemos ido desarrollando, para acabar podemos concluir –con las matizaciones que hemos ido apuntando– lo siguiente:

- La gran mayoría de las pruebas de las dos primeras cohortes se pueden considerar esencialmente unidimensionales o de débil multidimensionalidad.
- Por el contrario, las pruebas de la tercera cohorte son mayoritariamente de moderada multidimensionalidad (en términos relativos).
- La mayoría de las pruebas tienen una estructura de una cierta complejidad y las que resultan ser de estructura más simple tienden a mayor multidimensionalidad.
- Esto nos reafirma en la idea de que, mayoritariamente, nos encontramos con una dimensión latente principal de la que participan la gran mayoría de los ítems y una secundaria.
- En las pruebas de estructura simple, podría hablarse probablemente de dos dimensiones más marcadas. En la primera seguirían cargando muchos ítems, y la segunda, más claramente definida, se aglutinaría alrededor de contenidos como la geometría, la medida o la estadística o por dominios cognitivos de aplicación o resolución de problemas.
- Para las tres cohortes y desde la perspectiva de la dimensionalidad, las formas A y B no resultan equivalentes.
- Desde una perspectiva longitudinal, en las dos primeras cohortes la estructura dimensional de las pruebas se mantiene básicamente a lo largo de las cuatro aplicaciones. No ocurre lo mismo con la tercera.
- La triangulación realizada mediante los procedimientos factoriales «clásicos», el análisis factorial de información completa y los procedimientos no paramétricos basados en TRI ha resultado una estrategia metodológica válida que nos ha permitido acumular evidencia empírica que consideramos suficiente para apoyar las conclusiones anteriores.

Todos estos resultados confirman que el aumento de la complejidad de los contenidos implica un aumento en la complejidad de la estructura dimensional de las pruebas. Los modelos de medida utilizados son ciertamente bastante robustos a las violaciones leves de los supuestos relacionados con la independencia local.

En el futuro, siempre que la evaluación de valor añadido se refiera a un abanico de cursos tan amplio como el que aquí se ha estudiado, se recomienda tener muy en

cuenta la posibilidad de contar con estructuras de dimensionalidad de complejidad creciente en paralelo con el desarrollo del currículo.

## Referencias bibliográficas

- ABAD, F.J., PONSODA, V. Y REVUELTA, J. (2006). *Modelos politómicos de respuesta al ítem*. Madrid: La Muralla.
- FINCH, H. (2006). Comparison of the Performance of Varimax and Promax Rotations: Factor Structure Recovery for Dichotomous Ítems. *Journal of Educational Measurement*, 43, 39-52.
- GAVIRIA, J. L. Y RUIZ DE MIGUEL, C. (2007). Importancia de algunos supuestos psicométricos en la evaluación de los sistemas educativos. Calibración y equiparación en las pruebas de Estándares Nacionales de México. *Revista de Educación*, 343, 223-248.
- HATTIE, J., KRAKOWSKI, K., ROGERS H.J. & SWAMINATHAN, H. (1996). An Assessment of Stout's Index of Essential Unidimensionality. *Applied Psychological Measurement*. 20, 1-14.
- JANG, E.E. & ROUSSOS, L. (2007). An Investigation into the Dimensionality of TOEFL Using Conditional Covariance-Based Nonparametric Approach. *Journal of Educational Measurement*, 44, 1, 1-21.
- MARTÍNEZ, M. R., HERNÁNDEZ, M. J. Y HERNÁNDEZ, M. V. (2006). *Psicometría*. Madrid: Alianza.
- MONAHAN, P.O., STUMP, T.E., FINCH, H. & HAMBLETON, R.K. (2007). Bias of Exploratory and Cross-Validated DETECT Index Under Unidimensionality. *Applied Psychological Measurement*, 31, 483-503.
- NANDAKUMAR, R. & STOUT, W. (1993). Refinements of Stout's procedure for assessing unidimensionality. *Journal of Educational Statistics*. 18, 41-68.
- ROUSSOS, L.A. & OZBEK, O.Y. (2006). Formulation of the DETECT Population Parameter and Evaluation of DETECT Estimator Bias. *Journal of Educational Measurement*, 43, 3, 215-243.
- ROUSSOS, L.A. & STOUT, W. (2007). *Dimpack. Version 1.0*. The Roussos-Stout Software Development Group. St. Paul, MN: Assesment Systems Corp.
- STOUT, W.F. (1987). A nonparametric approach for assessing latent trait dimensionality. *Psychometrika*, 52, 589-617.
- (1990). A new item response theory modelling approach with applications to unidimensional assessment and ability estimation. *Psychometrika*, 55, 293-326.

- STOUT, W., HABING, B., DOUGLAS, J., KIM, H., ROUSSOS, L. & ZHANG, J. (1996). Conditional Covariance-Based Nonparametric Multidimensionality Assessment. *Applied Psychological Measurement*, 20, 331-354.
- STOUT, W., NANDAKUMAR, R. & HABING, R. (1996) Analysis Of Latent Dimensionality Of Dichotomously And Polytomously Scored Test Data. *Behaviormetrika*, 23, 1, 37-65.
- STOUT, W., FROELICH, A. G. & GAO, F. (2001). Using resampling methods to produce an improved DIMTEST procedure. En A. BOOMSMA, M.A.J. DUIJN, & T.A.B. SNIDJERS (Eds.), *Essays on item response theory* (357-376). NY: Springer-Verlag.
- WALKER, C.M., AZEN, R. & SCHMITT, T. (2006). Statistical Versus Substantive Dimensionality. The Effect of Distributional Differences on Dimensionality Assessment Using DIMTEST. *Educational and Psychological Measurement*. 66, 5, 721-738.
- WILSON, D.T., WORD, R. & GIBBONS, R. (1998). *TESTFACT. Test scoring, Item Statistics and Item Factor Analysis*. Chicago: Scientific Software Internacional, Inc.
- ZHANG, J. & STOUT, W. (1999). The theoretical Detect index of dimensionality and its application to approximate simple structure. *Psychometrika*, 64, 2, 213-249.
- ZHANG, J. (2007). Conditional Covariance Theory And Detect For Polytomous Items. *Psychometrika*, 72, 1, 69-91.

## Fuentes electrónicas

- PERKHOUNKOVA, Y. & DUNBAR, S. B. (1999). Influences of Item Content and Format on the Dimensionality of Tests Combining Multiple-Choice and Open-response ítems: An Application of the Poly-DIMTEST Procedure. Trabajo presentado en Annual Meeting of the American Educational Research Association, Montreal, Canadá, Abril. Consultado el 2 de febrero de 2008, de ERIC [www.eric.ed.gov](http://www.eric.ed.gov).
- WANG, Y & HOCEVAR, D. (1994). Effects of mathematics Test Content Specificity on Essential Dimensionality in U.S. and Japan Data. Trabajo presentado en *Annual Meeting of the American Educational Research Association*, New Orleans, LA, Abril, 4-8. Consultado el 2 de febrero de 2008, de ERIC [www.eric.ed.gov](http://www.eric.ed.gov)

**Dirección de contacto:** Luis Lizasoain Hernández. Universidad del País Vasco-Euskal Herriko Unibertsitatea. Avenida de Tolosa, 70. 20018 Donostia - San Sebastián. España. Departamento de Métodos de Investigación y Diagnóstico en Educación. E-mail: [luis.lizasoain@ehu.es](mailto:luis.lizasoain@ehu.es).

# Patrones de correlación entre medidas de rendimiento escolar en evaluaciones longitudinales: un estudio de simulación desde un enfoque multinivel

## Achievement measurements correlation patterns in longitudinal assessments: a multilevel approach simulation study

Ángeles Blanco Blanco, Coral González Barberá y Xavier G. Ordóñez

*Universidad Complutense. Facultad de Educación. Departamento de Métodos de Investigación y Diagnóstico en Educación (MIDE). Madrid, España.*

### Resumen

Los modelos de valor añadido operan sobre una estructura de datos longitudinal que plantea importantes retos metodológicos, tanto estadísticos como psicométricos. Este artículo estudia un problema particular asociado a esta estructura desde el enfoque de los modelos jerárquico-lineales con medidas repetidas. Concretamente se estudian los patrones diferenciales de correlación que presentan las puntuaciones latentes, observadas y verdaderas a medida que transcurre el tiempo, y el efecto que ejercen sobre ellos dos factores: la complejidad creciente del rasgo y los patrones de covariación de los residuos en el nivel del alumno y del centro. Para ello, se realiza una simulación que trabaja con una muestra de 25.000 sujetos, para los que se generan puntuaciones latentes, observadas y verdaderas en cuatro evaluaciones sucesivas y seis condiciones distintas. Los resultados permiten describir un patrón de correlaciones en las puntuaciones latentes cuando se asume la unidimensionalidad del rasgo, que difiere del registrado habitualmente para las puntuaciones observadas. Así, la magnitud de la correlación entre dos medidas es mayor cuanto más se alejan del punto inicial de la serie de medidas. Asimismo, los resultados permiten interpretar este patrón mediante el análisis comparativo del efecto de

los dos factores considerados sobre los distintos tipos de puntuaciones. Entre las conclusiones del estudio, destacamos dos con especial relevancia en la evaluación longitudinal del rendimiento académico: la necesidad de conocer los patrones de correlación en el nivel de las variables latentes para dar explicaciones plausibles a los patrones empíricos encontrados y la relevancia de la complejidad creciente del rasgo evaluado en la determinación de las estructuras de covariación.

*Palabras clave:* medida del rendimiento académico, análisis de datos longitudinales, evaluación del valor añadido, modelos multinivel para el estudio del cambio, teoría de respuesta al ítem, multidimensionalidad del rasgo latente, simulación.

### **Abstract**

Value-added models are applied in longitudinal data structures that raise important statistical and psychometrical methodological challenges. In this context, the aim of the study was to explain the differential correlation patterns of *observed*, *true* and *latent* scores over time. In particular, the effect of the *trait complexity factor* and the second (student) and third (school) levels of the *residual covariate patterns* was analyzed, from the point of view of the lineal hierarchical models with repeated measurements. *Observed*, *true* and *latent* scores for 25000 subjects were generated using a simulation, in four consecutive assessments and six different conditions. The first result is a description of the correlation pattern in the latent scores when the unidimensionality of the trait is assumed. This pattern is different to the one usually observed in empirical studies. Thus, the magnitude of the correlation between two measures increases as the measures distance to the starting point increases. This pattern can be explained as the result of the effect of the two referred factors. From the results two main findings are proposed. First, the need of understanding of the correlation patterns in the latent variables level to explain the relations in the empirical level. Second, the correlation patterns of achievement measures in educational longitudinal assessment are mainly determined by the growing complexity of the trait.

*Key Words:* achievement measure, testing, longitudinal data analysis, value-added assessment, multilevel models for change, item response theory, latent variable multidimensionality, simulation study.

## Introducción

Bajo la denominación genérica de *modelos de valor añadido* se incluyen en la actualidad diferentes modelos estadísticos aplicados a la evaluación de sistemas educativos que varían muy sustancialmente entre ellos, tanto en complejidad como en los supuestos que subyacen a los mismos (Sanders, 2006; Wiley, 2005). Sin embargo, todas las aproximaciones y propuestas englobadas bajo este término comparten un objetivo: vincular los cambios registrados en el rendimiento individual de los alumnos con las escuelas a las que asisten o con los profesores responsables de la clase a la que pertenecen (Martínez Arias, Gaviria y Castro, 2009). Puesto que son modelos orientados al análisis del cambio, un segundo denominador común a todos ellos, derivado del anterior, es que han de operar sobre datos que permitan el seguimiento individual del crecimiento en el rendimiento a lo largo del tiempo, con el fin de estimar la contribución de la escuela y/o el profesor a dicho crecimiento (Braun, 2005). En consecuencia, la estructura longitudinal de las medidas de rendimiento del alumnado y las implicaciones que de ésta se derivan conforman un objeto de análisis de importancia central en las aproximaciones metodológicas a la estimación de medidas del valor añadido. Porque ciertamente la modelización de datos longitudinales plantea un número importante de cuestiones estadísticas y psicométricas (para una revisión amplia, véase por ejemplo McCaffrey, Lakewood, Koretz y Hamilton, 2003; y, en este mismo número, Martínez Arias, 2009).

En el marco general de estas consideraciones, el presente trabajo se ocupa de analizar un problema particular ligado a la naturaleza longitudinal de las medidas del rendimiento académico típicamente usadas en la evaluación de sistemas educativos. En concreto, el objetivo central del trabajo es ilustrar y explicar la atenuación que se registra en las correlaciones halladas a lo largo del tiempo entre las puntuaciones de rasgo ( $\theta$ ), entre las puntuaciones observadas ( $X$ ) y entre las puntuaciones verdaderas ( $T$ ) en las evaluaciones longitudinales. Como se verá, este fenómeno puede entenderse presumiblemente asociado al incremento de la complejidad del constructo *rendimiento académico* en un área determinada a lo largo del tiempo, y por tanto asociado a la desviación de la unidimensionalidad supuesta del rasgo cuando se inicia la serie de medidas.

De acuerdo con lo que constituyen las propuestas recientes más aceptadas, el estudio adopta la aproximación proporcionada por los modelos jerárquico-lineales para la modelización longitudinal (Singer y Willet, 2003), alineándose igualmente con los desarrollos metodológicos más actuales en la medida del valor añadido. Como una derivación de la decisión anterior, también se incorpora como factor explicativo del

fenómeno bajo estudio el patrón diverso de covariación (positivo o nulo) asociado a los residuos en el nivel del alumno y del centro educativo.

## Planteamiento del problema

En el contexto de las evaluaciones longitudinales del rendimiento, se asume que la magnitud de la correlación observada entre dos medidas dadas (y por tanto la capacidad predictiva de una medida previa sobre una posterior) es tanto mayor cuanto mayor es también la proximidad temporal entre las mismas. De este modo, se espera que, en un esquema de evaluación con cuatro momentos de medida, la correlación entre la primera y la segunda medida sea superior a la que se registre entre la primera y la tercera, correlación que se espera a su vez superior a la correspondiente a la primera y la cuarta medida. Igualmente se asume que la magnitud de las correlaciones con idéntica proximidad temporal será básicamente equiparable bajo condiciones de fiabilidad constante, de modo que la correlación entre la primera y la segunda medida será similar a la registrada entre la tercera y la cuarta.

Sin embargo, la asunción de este patrón frecuente de covariación presumido para las puntuaciones observadas, ¿puede mantenerse cuando tomamos como referente las puntuaciones de rasgo? Es decir, ¿es el que cabe esperar cuando el análisis está referido a las relaciones *reales* que acontecen entre las variables latentes? La respuesta negativa a esta cuestión y sus implicaciones constituyen el centro de este trabajo.

Es importante considerar que hacemos aquí uso del término *real* en el sentido propuesto por Bhaskar (1978) cuando distingue entre tres dominios relevantes en la formalización científica: a) el *empírico*, que consiste en las impresiones de los sentidos y nuestras experiencias directas de las cosas; b) el dominio de *lo actual* que consiste en los eventos que se producen y dan lugar a la realidad efectiva, con toda su complejidad, tanto si es directamente experimentada por nosotros como si no; c) el dominio de *lo real*, que consiste en entidades y estructuras que producen eventos.

Así, desde una perspectiva postpositivista y en un contexto de realismo científico, no son los eventos en sí mismos los que constituyen la realidad, sino que lo actual es una de las posibles manifestaciones de lo real. Por lo tanto, como señala House (1991), desde una perspectiva realista las leyes que se formulan en la ciencia no son proposiciones acerca de los eventos observados o las experiencias habidas, sino más

bien proposiciones acerca de las formas en que las entidades causales actúan, produciendo de esa forma eventos y experiencias.

En definitiva, es el dominio de *lo real* el que debe ser conocido y explicado. Por tanto, es de la máxima relevancia el estudio y determinación de las relaciones causales que se establecen entre las variables latentes, puesto que lo que se hace empírico y lo que se hace observable (los patrones correlacionales registrados en las variables observadas) depende de los patrones de correlación que operan a ese nivel.

Considerado lo anterior, el planteamiento del problema parte de un modelo jerárquico lineal básico para la determinación de puntuaciones de rasgo o latentes en el constructo rendimiento, con tres niveles y en el que se incorpora como predictor la variable tiempo:

$$\begin{array}{ll}
 \blacksquare \text{ Primer nivel: Ocasión/Tiempo} & y_{ijk} = \beta_{ojk} + \beta_{1jk}(t - t_o) + \varepsilon_{ijk} \\
 \blacksquare \text{ Segundo nivel: Alumno} & \beta_{ojk} = \beta_{ok} + \mu_{ojk} \quad \beta_{1jk} = \beta_{1k} + \mu_{1jk} \\
 \blacksquare \text{ Tercer nivel: Escuela} & \beta_{ok} = \beta_{oo} + \mu_{ok} \quad \beta_{1k} = \beta_{10} + \mu_{1k} \quad (0)
 \end{array}$$

Un hecho derivado directamente del modelo mismo es que la varianza total de las puntuaciones en rendimiento aumenta a medida que lo hace el tiempo. Sin embargo, es muy importante notar que, a la vez, el modelo asume que el peso relativo del término residual  $\varepsilon_{ijk}$  permanece constante a lo largo del tiempo. En consecuencia, es la varianza sistemática la que aumenta manteniéndose constante la varianza no explicada, lo que quiere decir que la varianza compartida entre las variables de respuesta correspondientes a las distintas aplicaciones (cuantificada en términos de covarianza/correlación) aumenta progresivamente. Entonces, es claro que en estas condiciones cabe esperar que las correlaciones entre aplicaciones aumenten a medida que transcurre el tiempo y se van sumando evaluaciones o medidas consecutivas. Lo anterior, que puede ser interpretado en términos de aumento creciente del poder explicativo del modelo, alejará el patrón correlacional registrado para las puntuaciones de rasgo del patrón generalmente asumido para las correlaciones observadas entre evaluaciones sucesivas de rendimiento.

## Una ejemplificación de patrones de correlación entre puntuaciones de rasgo

Con el fin de ilustrar la proposición anterior, se desarrolla un ejemplo en el que se atribuyen valores a los términos del modelo multinivel básico [0]. De este modo, se obtendrán las correlaciones para cuatro aplicaciones supuestas de un rasgo dado.

La correlación entre dos medidas del rasgo y tomadas en los momentos t y s puede definirse como:  $r_{y_t, y_s} = \sigma_{y_t, y_s} / \sigma_{y_t} \sigma_{y_s}$

Y de acuerdo con el modelo [0], la medida del rasgo y en los momentos t y s toma la forma siguiente:

$$y_t = \beta_{0jk} + \beta_{1jk}(t - t_0) + \varepsilon_{tjk} ;$$

$$y_s = \beta_{0jk} + \beta_{1jk}(s - t_0) + \varepsilon_{sjk}$$

siendo por tanto la varianza de la medida, ilustrada aquí para el momento t:

$$\begin{aligned} \sigma_{y_t}^2 = & \sigma_{\mu_{0k}}^2 + \sigma_{\mu_{0jk}}^2 + (t - t_0)^2 (\sigma_{\mu_{1k}}^2) + (t - t_0)^2 (\sigma_{\mu_{1jk}}^2) + 2(t - t_0) (\sigma_{\mu_{0k}\mu_{1k}}) + \\ & + 2(t - t_0) (\sigma_{\mu_{0jk}\mu_{1jk}}) + \sigma_{\varepsilon_{tjk}}^2 \end{aligned} \quad (1)$$

La covarianza entre medidas del rasgo en los tiempos t y s, por su parte, es igual a:

$$\sigma_{y_t, y_s} = E(y_t y_s) - E(y_t)E(y_s)$$

Donde:

$$E(y_t) = E(\beta_{0jk}) + (t - t_0)E(\beta_{1jk}) = E(\beta_{0k}) + (t - t_0)E(\beta_{1k}) = \beta_{00} + (t - t_0)\beta_{10}$$

; siendo por tanto  $E(y_s) = \beta_{00} + (s - t_0)\beta_{10}$

Así tenemos que:

$$\begin{aligned} E(y_t)E(y_s) = & \beta_{00}^2 + (s - t_0)\beta_{10}\beta_{00} + (t - t_0)\beta_{10}\beta_{00} + \beta_{10}^2(t - t_0)(s - t_0) = \\ = & \beta_{00}^2 + \beta_{10}\beta_{00}[(t - t_0) + (s - t_0)] + \beta_{10}^2[(t - t_0)(s - t_0)] \end{aligned}$$

Por otro lado,

$$\begin{aligned} E(y_t y_s) = & E[\beta_{0jk}^2 + \beta_{0jk}\beta_{1jk}[(t - t_0) + (s - t_0)] + \beta_{1jk}[(t - t_0)(s - t_0)]] = \\ = & E(\beta_{0jk}^2) + [(t - t_0) + (s - t_0)]E(\beta_{0jk}\beta_{1jk}) + E(\beta_{1jk}^2) \end{aligned}$$

Donde:

$$\begin{aligned} E(\beta_{0jk}^2) = & E[(\beta_{00} + \mu_{0k} + \mu_{0jk})^2] = E[\beta_{0k}^2 + \mu_{0jk}^2 + 2\beta_{0k}\mu_{0jk}] = \\ = & E[\beta_{00}^2] + E[\mu_{0k}^2] + E[\mu_{0jk}^2] ; \end{aligned}$$

$$E(\beta_{1jk}^2) = E[\beta_{10}^2] + E[\mu_{1k}^2] + E[\mu_{1jk}^2] ; y$$

$$E(\beta_{0jk}\beta_{1jk}) = E[\beta_{0k}\beta_{1k} + \beta_{0k}\beta_{1jk} + \beta_{1k}\beta_{0jk} + \mu_{0jk}\mu_{1jk}] ;$$

Puesto que  $E(\beta_{ok}\beta_{1jk}) = E(\beta_{1k}\beta_{ojk}) = 0$ , tenemos que:

$$E(\beta_{ojk}\beta_{1jk}) = E[\beta_{ok}\beta_{1k}] + E[\mu_{ojk}\mu_{1jk}] = \beta_{00}\beta_{10} + E[\mu_{0k}\mu_{1k}] + E[\mu_{0jk}\mu_{1jk}]$$

Y sustituyendo los tres términos anteriores en la expresión inicial:

$$E(y_t y_s) = \beta_{00}^2 + E[\mu_{0k}^2] + E[\mu_{0jk}^2] + [(t-t_0) + (s-t_0)]\beta_{00}\beta_{10} + E(\mu_{0k}\mu_{1k}) + E(\mu_{0jk}\mu_{1jk}) + [(t-t_0)(s-t_0)] [E[\mu_{1k}^2] + E[\mu_{1jk}^2]]$$

de donde:

$$\begin{aligned} \sigma(y_t, y_s) &= \beta_{00}^2 + [(t-t_0) + (s-t_0)]\beta_{00}\beta_{10} + (t-t_0)(s-t_0)\beta_{10}^2 - \beta_{00}^2 - \\ &- [(t-t_0) + (s-t_0)]\beta_{00}\beta_{10} - (t-t_0)(s-t_0)\beta_{10}^2 + E[\mu_{0k}^2] + E[\mu_{0jk}^2] + \\ &+ [(t-t_0) + (s-t_0)] [E(\mu_{0k}\mu_{1k}) + E(\mu_{0jk}\mu_{1jk})] + (t-t_0)(s-t_0)\mu_{1k}^2 + \\ &+ (t-t_0)(s-t_0)\mu_{1jk}^2 \end{aligned}$$

Simplificando la expresión anterior, tenemos finalmente que la covarianza entre las puntuaciones de dos aplicaciones consecutivas puede expresarse como sigue:

$$\begin{aligned} \sigma(y_t, y_s) &= \sigma^2(\mu_{0k}) + \sigma(\mu_{0k}\mu_{1k}) [(t-t_0) + (s-t_0)] + \\ &+ \sigma^2(\mu_{1k}) [(t-t_0)(s-t_0)] + \sigma^2(\mu_{0jk}) + \sigma(\mu_{0jk}\mu_{1jk}) [(t-t_0) + (s-t_0)] + \\ &+ \sigma^2(\mu_{1jk}) [(t-t_0)(s-t_0)] \end{aligned} \quad (2)$$

Este último resultado analítico es de la máxima importancia. Puesto que la covarianza entre dos medidas dadas del rasgo está fuertemente determinada por la distancia que las separa temporalmente del momento original de medida ( $t_0$ ), supuesta una covariación no nula de los residuos, la covariación de las puntuaciones se incrementa progresivamente de modo sustancial como un simple resultado del paso del tiempo. Este hecho no es coherente con la experiencia habitual, en la que lo que realmente se observa es una atenuación de las correlaciones.

Con el fin de mantener la ilustración en los términos más sencillos posibles, se adopta una métrica familiar que fija todas las varianzas a 1, valor arbitrario considerado útil a los efectos descritos. Los valores asignados a las covarianzas tratan de reproducir además un contexto plausible en el escenario general de la evaluación del rendimiento. De este modo, se contempla tanto la existencia de una relación

sustancial entre los puntos de partida y la tasa de crecimiento registrada al nivel de escuela y de alumno como la independencia entre ambas características, bien sea parcial o totalmente. De este modo, para las covarianzas de los residuos tanto del nivel alumno ( $\sigma_{\mu ojk, \mu 1jk}$ ) como de centro ( $\sigma_{\mu ok, \mu 1k}$ ) se toman dos valores alternativos iguales, a 0,7 y 0.

Como resultado de tales decisiones, las condiciones o escenarios considerados en el ejemplo son entonces los mostrados en la Tabla I.

TABLA I. Escenarios definidos para la reproducción de patrones de correlación entre medidas sucesivas del rasgo (rendimiento)

Escenario	$\sigma^2_{\mu 0k}$	$\sigma^2_{\mu 1k}$	$\sigma^2_{\mu 0jk}$	$\sigma^2_{\mu 1jk} \cdot \epsilon_{tjk}$	$\sigma_{\mu 0k, \mu 1k}$	$\sigma_{\mu 0jk, \mu 1jk}$
1						0,7
2						0,0
3						0,0

Para hallar la correlación entre el momento 0 y el momento 1 en el escenario 1, se procede entonces como sigue:

- En primer lugar, obtenemos la covarianza entre las puntuaciones sustituyendo en [2] los valores correspondientes a la varianza de los residuos y a sus covarianzas. Además,  $t$  y  $t_0$  son iguales en este caso a 0 y  $s$  toma en este ejemplo valor 1. Tenemos entonces que:

$$\sigma_{(y_0, y_1)} = 1 + 0,7[(0 - 0) + (1 - 0)] + 1[(0 - 0)(1 - 0)] + 1 + 0,7[(0 - 0) - (1 - 0)] + 1[(0 - 0)(1 - 0)] = 3,4$$

- A continuación hallamos los valores correspondientes a las varianzas de  $y_0$  e

$$\sigma^2_{y_0} = 1 + 1 + (0 - 0)^2(1) + (0 - 0)^2(1) + 2(0 - 0)(0,7) + 2(0 - 0)(0,7) + 1 = 3$$

$$\sigma^2_{y_1} = 1 + 1 + (1 - 0)^2(1) + (1 - 0)^2(1) + 2(1 - 0)(0,7) + 2(1 - 0)(0,7) + 1 = 7,8$$

- Entonces podemos establecer que:

$$r_{y_0, y_1} = \sigma_{y_0, y_1} / \sigma_{y_0} \sigma_{y_1} = \frac{3,4}{\sqrt{3} \sqrt{7,8}} = 0,703$$

Generalizando el procedimiento para los cuatro momentos de medida y los tres escenarios definidos, se obtienen los resultados que se muestran en la Tabla II.

TABLA II. Correlaciones entre medidas sucesivas del rasgo derivadas de las varianzas/covarianzas de los residuos del modelo jerárquico-lineal especificado

	Escenario 1	Escenario 2	Escenario 3
$r_{y0y1}$	0,703	0,616	0,516
$r_{y0y2}$	0,680	0,528	0,348
$r_{y0y3}$	0,660	0,472	0,252
$r_{y1y2}$	0,896	0,862	0,809
$r_{y1y3}$	0,898	0,850	0,781
$r_{y2y3}$	0,951	0,938	0,921

Los resultados muestran que efectivamente el patrón esperado se reproduce en los tres escenarios.

En primer lugar, se constata que inicialmente la correlación entre aplicaciones consecutivas es ligeramente mayor que entre aplicaciones distanciadas en el tiempo, de modo que entre los momentos 0 y 1 la magnitud de la correlación es superior a la registrada entre los momentos 0 y 2 ó 0 y 3. Pero lo que especialmente cabe destacar es que a medida que pasa el tiempo la magnitud de las correlaciones encontradas es mayor, tanto entre aplicaciones sucesivas como entre aplicaciones distanciadas. Esto es, entre los momentos 0 y 1 la correlación es sensiblemente inferior a la encontrada entre los momentos 2 y 3, a pesar de existir la misma *distancia temporal* entre unos y otros. Y otro tanto ocurre si se comparan las correlaciones entre los momentos 0 y 2 con las correspondientes a los momentos 1 y 3. La explicación a este fenómeno, como ya se apuntó, puede hacerse en términos de «varianza disponible», puesto que, si se mantiene el error constante (característica común de los tres escenarios), su importancia relativa decrece a lo largo del tiempo, aumentando la proporción de varianza sistemática y por tanto la magnitud de las correlaciones entre aplicaciones.

Complementariamente, es importante subrayar que el valor absoluto de las correlaciones, manteniéndose constante el patrón descrito, desciende ligeramente al pasar del escenario 1 al 2 y de este último al 3. Es decir, cuando la covarianza de los residuos en los niveles de alumno y escuela es alta y positiva, la covariación entre las puntuaciones de rasgo latente es mayor que cuando los residuos de los dos niveles presentan parcial o totalmente correlaciones nulas.

## Preguntas de investigación e hipótesis

Las preguntas de investigación a las que se trata de dar respuesta en este trabajo se formulan en los siguientes términos:

- El patrón de correlaciones que frecuentemente observamos en las evaluaciones longitudinales del rendimiento no parece ajustarse a lo que cabe esperar en el nivel en el que acontecen las relaciones causales reales (entre variables latentes). Entonces: ¿se puede explicar dicha desviación como un efecto de la creciente complejidad del rasgo objeto de medida y por tanto de la creciente introducción de incertidumbre o varianza no explicada por el modelo a medida que transcurre el tiempo y se suman evaluaciones sucesivas del rendimiento?
- ¿Cómo afecta en este contexto, cuando se hace uso de modelos multinivel, la naturaleza de la covariación en los residuos al nivel de centro educativo y alumno?

En la práctica de la evaluación del rendimiento, aunque es claro que se pretende medir un único e idéntico rasgo a lo largo del tiempo (rendimiento en una materia o área de competencia), ciertamente ello resulta difícil. Así, basta pensar que en cada aplicación o momento de evaluación adaptamos la prueba a los contenidos del currículo (al que progresivamente se incorporan nuevos aspectos y facetas) y a la capacidad del alumno (presumiblemente creciente). Por tanto, parece razonable asumir que si bien se parte de un rasgo único en la primera aplicación, que nunca deja de ser el principal, al mismo van uniéndose otros rasgos o dimensiones que entran en juego a medida que avanzamos en el tiempo, alejándose así de la unidimensionalidad el objeto de la medida. Y, puesto que conforme pasa el tiempo el rasgo que queremos medir se hace más complejo, el modelo utilizado pierde en igual medida capacidad para determinar el mismo.

Esta última consideración conforma la hipótesis del trabajo: se espera que la simulación propuesta permita identificar la suavización de los patrones crecientes de correlación entre aplicaciones a medida que transcurre el tiempo cuando se considera la complejidad creciente del rasgo.

## Método

### Diseño general de la simulación

La simulación adopta el diseño general de un trabajo de evaluación llevado a cabo en la Comunidad Autónoma de Madrid<sup>1</sup>. En dicho estudio se tomaron medidas de rendimiento a una cohorte de alumnos en el inicio y al término de dos cursos académicos consecutivos.

Para la modelización del rendimiento se define un modelo jerárquico lineal con tres niveles. El modelo incluye como variable predictora, además del tiempo o momento de la medida, el tipo de centro educativo, con cuatro valores distintos. Se trata así de reproducir un contexto plausible para la simulación que atiende al hecho de que el alumnado no se distribuye aleatoriamente entre las distintas escuelas y que el rendimiento medio de éstas, por tanto, varía en cualquier escenario educativo real.

Cuando en el modelo general [0] se incorpora el predictor tipo de centro ( $W_k$ ) en el nivel macro correspondiente a la escuela, tenemos:

$$\text{Nivel escuela: } \beta_{0k} = \beta_{00} + \beta_{01}W_K + \mu_{0k} \quad \beta_{1k} = \beta_{10} + \beta_{11}W_K + \mu_{1k}$$

Por lo que, sustituyendo en los niveles anteriores y ordenando la parte fija y aleatoria, el modelo final en el nivel 1 toma la forma siguiente:

$$y_{tjk} = \beta_{00} + \beta_{01}W_K + \beta_{10}(t - t_0) + \beta_{11}W_k(t - t_0) + \mu_{0k} + \mu_{0jk} + \quad (3) \\ + \mu_{1k}(t - t_0) + \mu_{1jk}(t - t_0) + \varepsilon_{tjk}$$

donde  $t-t_0$  y  $W_k$  toman valores entre 0 y 3; es decir, contamos con cuatro momentos de medida y cuatro tipos de escuela diferenciadas en razón de su rendimiento.

Pues bien, la complejidad creciente asociada al rasgo en perspectiva longitudinal puede ser incluida en el modelo [3] multiplicando el término  $\varepsilon_{tjk}$  (residuo del individuo) por la variable tiempo, de modo que *de facto* se aumenta la importancia o relevancia de posibles dimensiones diferentes de la única e inicialmente considerada. Conviene destacar que la incorporación de este término es un modo de introducir la multidimensionalidad del rasgo en el modelo sin que se prejuzgue ni el

<sup>(1)</sup> El valor añadido en educación y la función de producción educativa: un estudio longitudinal; I+D SEC2003-09742.

número ni la naturaleza de las dimensiones involucradas en el mismo. También es claro que, si bien la complejidad creciente del rasgo no puede ser vista como la única fuente de incremento progresivo del error, no es menos cierto que en el contexto de la medida longitudinal del rendimiento escolar constituye la interpretación más plausible. Y en todo caso, al aumento de aquélla se asocia necesariamente el incremento de éste.

El modelo resultante, que presenta la introducción de un término residual modificado, quedaría expresado como sigue:

$$y_{ijk} = \beta_{00} + \beta_{01}W_K + \beta_{10}(t - t_0) + \beta_{11}W_k(t - t_0) + \mu_{0k} + \mu_{ojk} + \mu_{1k}(t - t_0) + \mu_{1jk}(t - t_0) + \varepsilon_{ijk}(t) \quad (4)$$

En consecuencia, se puede definir la capacidad de predicción o determinación del modelo sobre la base de lo que aquí denominamos *factor de complejidad* (FC). Dicho factor expresa la proporción de varianza que un rasgo único es capaz de explicar del total de la varianza registrada:

$$FC = \frac{\sigma_{\theta_j}^2}{\sigma_y^2}$$

Si FC es igual o cercano a 1, la complejidad del rasgo es nula, es decir, el rasgo es unidimensional y por tanto la importancia relativa del residuo del individuo en el modelo se mantendría constante. Por el contrario, si FC es menor que 1, la complejidad aumenta a medida que el valor de FC es más pequeño, lo que indica la multidimensionalidad del rasgo y tiene su reflejo en el aumento de la importancia relativa del residuo del individuo con el paso del tiempo.

La simulación, por tanto, considera dos modelos para generar puntuaciones de rasgo. A partir de cada uno de ellos, se desarrollan tres escenarios que se distinguen por los valores de las covarianzas de los residuos, que fueron ya definidos y justificados en la ejemplificación de los patrones de correlación que sirvió para plantear el problema. Los tres primeros escenarios, derivados del modelo [3], asumen el carácter constante del residuo del individuo. El modelo [4] se emplea para generar puntuaciones en el contexto de los tres escenarios restantes, caracterizados por el carácter creciente del residuo  $\varepsilon_{ijk}$ . Por tanto, el esquema general de la simulación es el que muestra la Tabla III.

**TABLA III. Conjunto de escenarios considerados en la simulación**

FC=I				FC<I			
Escenario	$\sigma_{\mu_{0k}, \mu_{1k}}$	$\sigma_{\mu_{0k}, \mu_{2k}}$	$\epsilon_{ijk}$	Escenario	$\sigma_{\mu_{0k}, \mu_{1k}}$	$\sigma_{\mu_{0k}, \mu_{2k}}$	$\epsilon_{ijk}$
1	0,7	0,7	Constante	4	0,7	0,7	Crecente
2	0,7	0,0	Constante	5	0,7	0,0	Crecente
3	0,0	0,0	Constante	6	0,0	0,0	Crecente

FC: Factor de complejidad

### Muestra y generación inicial de puntuaciones en el rasgo latente

La muestra está conformada por 25.000 sujetos distribuidos en 1.000 escuelas, contándose con 25 sujetos por grupo y un solo grupo en cada escuela. Se definen cuatro tipos de escuelas atendiendo al nivel de rendimiento de su alumnado, siendo las escuelas de tipo I las de menor rendimiento medio y las de tipo IV las de rendimiento superior. Un tercio de la muestra se define como escuelas tipo II y otro tercio, de tipo III. El tercio de escuelas restante se reparte entre las de tipo I y IV.

Fijado lo anterior, las puntuaciones de rasgo se generaron directamente a partir de los modelos [3] y [4].

En lo que se refiere a las *variables predictoras*, ya se indicó anteriormente que  $t-t_0$  y  $W_k$  toman valores entre 0 y 3.

Los valores de los coeficientes fueron fijados atendiendo a dos criterios básicos: plausibilidad y sencillez de la escala métrica, determinada por una media igual a 0 y una desviación típica igual a 1. De este modo, se definieron los coeficientes que se describen a continuación:

- La constante o intercepto  $\beta_{00}$  se fija a 0
- Al coeficiente  $\beta_{01}$  se le asigna un valor igual a 0,4, contándose con una diferencia máxima razonable entre las escuelas de menor y mayor rendimiento (igual a 1,2).
- Tanto  $\beta_{10}$  como  $\beta_{11}$  han sido fijados en un valor de 0,2, de nuevo arbitrario pero perfectamente plausible en una escala de media 0 y desviación típica 1. De este modo entre la primera y la cuarta evaluación se considera una tasa total de crecimiento medio igual a 0,8, que es también la diferencia máxima en las tasas de cambio entre las escuelas (la registrada entre las escuelas tipo I y tipo IV).

Por lo que se refiere a la parte aleatoria del modelo, los residuos de los puntos de corte en los niveles alumno y escuela ( $\mu_{0ik}, \mu_{0jk}$ ) fueron generados directamente como

variables aleatorias distribuidas según el modelo  $N\sim(0,1)$ . A continuación, se introdujeron como variables independientes en la ecuación de regresión definida para obtener los residuos de las pendientes, de modo que:

$$\mu_{1k} = \mu_{0k} r_{\mu_{1k}, \mu_{0k}} + \varepsilon \sqrt{1 - r_{\mu_{1k}, \mu_{0k}}^2}, \text{ siendo } \varepsilon \sim N(0,1)$$

$$\mu_{1jk} = \mu_{0jk} r_{\mu_{1jk}, \mu_{0jk}} + \varepsilon \sqrt{1 - r_{\mu_{1jk}, \mu_{0jk}}^2}, \text{ siendo } \varepsilon \sim N(0,1)$$

Puesto que asumimos que las varianzas de los cuatro tipos de residuos toman valor 1, entonces:

$$r_{\mu_{1k}, \mu_{0k}} = \frac{\sigma_{\mu_{1k}, \mu_{0k}}}{\sigma_{\mu_{1k}} \sigma_{\mu_{0k}}}; r_{\mu_{1k}, \mu_{0k}} = \sigma_{\mu_{1k}, \mu_{0k}}, y$$

$$r_{\mu_{1jk}, \mu_{0jk}} = \frac{\sigma_{\mu_{1jk}, \mu_{0jk}}}{\sigma_{\mu_{1jk}} \sigma_{\mu_{0jk}}}; r_{\mu_{1jk}, \mu_{0jk}} = \sigma_{\mu_{1jk}, \mu_{0jk}}$$

Por tanto:

$$\mu_{1k} = \mu_{0k} \sigma_{\mu_{1k}, \mu_{0k}} + \varepsilon \sqrt{1 - \sigma_{\mu_{1k}, \mu_{0k}}^2}, \text{ siendo } \varepsilon \sim N(0,1)$$

$$\mu_{1jk} = \mu_{0jk} \sigma_{\mu_{1jk}, \mu_{0jk}} + \varepsilon \sqrt{1 - \sigma_{\mu_{1jk}, \mu_{0jk}}^2}, \text{ siendo } \varepsilon \sim N(0,1)$$

Las covarianzas de los residuos tanto del nivel alumno ( $\sigma_{\mu_{0jk}, \mu_{1jk}}$ ) como de centro ( $\sigma_{\mu_{0k}, \mu_{1k}}$ ) se fijan en dos valores alternativos (0,7 y 0) dependiendo de los escenarios de simulación considerados.

Finalmente, el término aleatorio del error asociado al sujeto ( $\varepsilon_{ijk}$ ) se generó en todos los escenarios de trabajo como distribuido normalmente, con media cero y varianza fijada a 1.

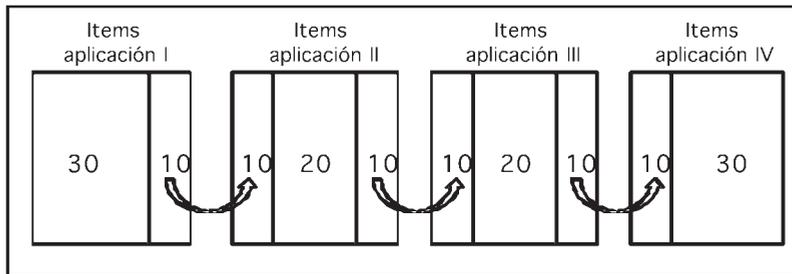
Pues bien, a partir de los modelos y de las especificaciones descritas, se obtuvieron con SPSS 15.0 las puntuaciones generadas iniciales de rasgo para cada sujeto y para cada aplicación en cada uno de los escenarios simulados para su comparación. Se generaron, por tanto, 6 matrices de 25.000 sujetos \* 4 puntuaciones de rasgo o latentes.

## Puntuaciones observadas y puntuaciones verdaderas

Para la simulación de las puntuaciones observadas o directas se siguió el proceso inverso al habitual, llevándose a cabo en el marco de la Teoría de Respuesta al Ítem y particularmente empleando un modelo logístico de un parámetro o modelo de Rasch (Rasch, 1960).

La calibración se realizó en el entorno de **R** (Ihaka y Gentleman, 1996) haciendo uso de la librería *irt*. Así, en primer lugar se generaron aleatoriamente valores plausibles del parámetro  $b$  para un conjunto de 120 ítems aplicados a lo largo de las cuatro aplicaciones simuladas. El esquema definido para cada uno de los tests, incluidas las secciones con ítems de anclaje (cuyos parámetros quedan fijos entre aplicaciones), puede verse en la Figura I, y se tomó de la investigación ya citada.

FIGURA I. Estructura de las pruebas simuladas e ítems de anclaje



La generación aleatoria del parámetro de dificultad se realizó conforme a una Distribución Normal (0,1) con puntuaciones entre -2 y +2, haciendo que aumentasen entre aplicaciones o momentos de medida de forma proporcional al rendimiento.

El parámetro definido en el paso anterior permite determinar la probabilidad de acierto a cada ítem dado el nivel de habilidad de cada sujeto en cada aplicación. Aplicando sobre dichas probabilidades los valores generados para una distribución uniforme, todo valor menor o igual que  $p$  pasa a ser definido como acierto en un ítem dado, mientras que, si dicho valor es mayor que  $p$ , el ítem correspondiente pasa a ser codificado como incorrectamente resuelto. Así, se obtiene la puntuación total observada ( $X$ ) de cada sujeto en cada aplicación y escenario a partir de la suma de aciertos de cada sujeto en cada uno de los 120 ítems.

La suma de las probabilidades de acierto a cada ítem permite la obtención de la puntuación verdadera ( $T$ ) para cada sujeto en cada aplicación.

## Resultados

El conjunto de correlaciones obtenidas para cada uno de los seis escenarios definidos en la simulación y para los tres tipos de puntuaciones consideradas se muestra en la Tabla IV.

El primer resultado de interés es el referido a la capacidad de la simulación para reproducir los patrones de correlación de las variables latentes tal y como fueron definidos y ejemplificados a partir del modelo multinivel [0]. Efectivamente, la inspección de las correlaciones obtenidas ahora mediante simulación a partir del modelo [3] en los tres primeros escenarios (en los que el residuo permanece constante) permite constatar que las diferencias son inapreciables cuando se comparan con aquéllas que se estimaron y presentaron en la Tabla II. La conclusión es, por tanto, clara: las correlaciones de las puntuaciones de rasgo generadas se reproducen como cabe esperar por el modelo, es decir, el funcionamiento del modelo [3] es correcto.

TABLA IV. Correlaciones obtenidas en la simulación para los distintos escenarios y tipos de puntuaciones a lo largo de las cuatro aplicaciones supuestas (N=25.000)

Puntuaciones de rasgo/latentes	Escenario <sup>a</sup>					
	1	2	3	4	5	6
$r_{y_0i}$	,711	,606	,539	,606	,506	,427
$r_{y_02}$	,687	,510	,389	,562	,421	,302
$r_{y_03}$	,666	,447	,297	,543	,366	,230
$r_{y_12}$	,900	,876	,823	,635	,603	,507
$r_{y_13}$	,899	,863	,797	,626	,594	,493
$r_{y_23}$	,951	,947	,925	,635	,634	,548
<b>Puntuaciones observadas</b>						
$r_{y_0i}$	,651	,551	,490	,549	,451	,374
$r_{y_02}$	,620	,449	,341	,491	,360	,254
$r_{y_03}$	,585	,384	,251	,460	,307	,184
$r_{y_12}$	,842	,819	,769	,558	,523	,438
$r_{y_13}$	,824	,788	,729	,535	,501	,418
$r_{y_23}$	,911	,904	,879	,540	,530	,454
<b>Puntuaciones verdaderas</b>						
$r_{y_0i}$	,674	,574	,509	,567	,469	,390
$r_{y_02}$	,638	,464	,351	,508	,371	,264
$r_{y_03}$	,601	,396	,257	,474	,317	,192
$r_{y_12}$	,859	,838	,790	,566	,533	,445
$r_{y_13}$	,839	,804	,747	,544	,508	,425
$r_{y_23}$	,922	,916	,892	,544	,536	,458

<sup>a</sup> Escenario 1:  $\sigma_{\mu_0k_1ijk}$  y  $\sigma_{\mu_0k_2ijk}$  altas y positivas y  $\epsilon_{ijk}$  constante. Escenario 2:  $\sigma_{\mu_0k_1ijk}$  alta y positiva,  $\sigma_{\mu_0k_2ijk}$  nula y  $\epsilon_{ijk}$  constante. Escenario 3:  $\sigma_{\mu_0k_1ijk}$  y  $\sigma_{\mu_0k_2ijk}$  nulas y  $\epsilon_{ijk}$  constante. Escenario 4:  $\sigma_{\mu_0k_1ijk}$  y  $\sigma_{\mu_0k_2ijk}$  altas y positivas y  $\epsilon_{ijk}$  creciente. Escenario 5:  $\sigma_{\mu_0k_1ijk}$  alta y positiva,  $\sigma_{\mu_0k_2ijk}$  nula y  $\epsilon_{ijk}$  creciente. Escenario 6:  $\sigma_{\mu_0k_1ijk}$  y  $\sigma_{\mu_0k_2ijk}$  nulas y  $\epsilon_{ijk}$  creciente.

Verificado lo anterior, corresponde analizar los resultados que permiten dar respuesta a la pregunta central de la investigación, a saber, la importancia de los efectos del factor de complejidad del rasgo latente así como de las covarianzas de los residuos del modelo en los niveles de alumno y de escuela, en los tres tipos de puntuaciones: de rasgo, observadas y verdaderas.

En primer lugar, conviene considerar las correlaciones de las *puntuaciones de rasgo latente* en los escenarios 4, 5 y 6, con el fin de verificar la hipótesis de su atenuación cuando se comparan con los escenarios 1, 2 y 3. La Tabla IV pone de manifiesto cómo efectivamente, aunque el patrón correlacional se reproduce en lo básico, se atenúan en gran medida tanto las magnitudes de las correlaciones en términos absolutos como las diferencias relativas entre las mismas, lo que apunta al carácter determinante de la única variable diferencial entre ambos bloques: el factor de complejidad. Parece confirmarse por tanto que el modelo utilizado va perdiendo capacidad relativa de determinación a medida que pasa el tiempo; o lo que es lo mismo, que el efecto causal disminuye, tal y como se hipotetizó.

En definitiva, las diferencias entre las correlaciones correspondientes a los escenarios que asumen error constante (1 a 3) y aquéllas derivadas de escenarios que lo asumen creciente (4 a 6) son sensiblemente superiores a las encontradas dentro de cada uno de estos dos bloques. De este resultado puede extraerse como fundamental consecuencia que la importancia o relevancia del factor de complejidad del rasgo latente al determinar las puntuaciones en el mismo parece ser sensiblemente mayor que la presentada por la covarianza entre residuos de ambos niveles. Es decir, como ya se apuntó cuando se analizó la descomposición de las varianzas y covarianzas de los residuos, se comprueba ahora que basta partir de una covarianza entre residuos de segundo y tercer nivel positiva aunque poco elevada o casi nula en la primera aplicación para que la varianza sistemática final, tras mediciones consecutivas, aumente considerablemente, viéndose fuertemente atenuada cuando incluimos un factor multiplicativo del error ligado al tiempo y asociado al efecto que pudieran tener otros rasgos asociados al que deseamos medir (caso de los tres últimos escenarios).

Si se centra el análisis en las *puntuaciones observadas*, las únicas disponibles en un estudio empírico, se verifica cómo los patrones en todos los escenarios son muy similares a los obtenidos en las variables latentes, pero más atenuados. La explicación lógicamente está asociada al error de estimación.

Finalmente, el patrón de correlaciones de las *puntuaciones verdaderas* es prácticamente idéntico al de las puntuaciones observadas. Las escasas diferencias encontradas, que apuntan a una ligera superioridad en la magnitud de las correlaciones en las puntuaciones verdaderas, se deben únicamente al efecto de los errores cometidos al medir.

En definitiva, el error de estimación creciente atenúa las diferencias en las correlaciones de los escenarios 4, 5, 6 convirtiendo en irrelevante el factor *momento de aplicación*, variable que determinaba, por el contrario, el patrón de correlaciones en los escenarios 1, 2 y 3. De hecho, para las puntuaciones observadas, en alguno de tales escenarios el error de la medida llega incluso a compensar por completo dicho factor, un efecto por tanto claramente relacionado con la fiabilidad del instrumento.

En resumen, las diferencias encontradas en los patrones de correlaciones de los diferentes escenarios, definidos en función del factor de complejidad del rasgo y de las covarianzas de residuos de los niveles dos y tres del modelo multinivel, se deben principalmente al citado factor, y sólo en segundo lugar, de manera más moderada, a las covarianzas de los residuos. Además, es interesante destacar las diferencias encontradas en la comparación de los tres tipos de puntuaciones en cualquiera de los escenarios. Así, se registra una moderación de las magnitudes de las correlaciones en las puntuaciones observadas y verdaderas debido a la influencia de las características del instrumento utilizado para medir, aunque permanece prácticamente constante el patrón que ya se daba en las puntuaciones de rasgo o latentes.

## Conclusiones, limitaciones y prospectiva

De la simulación presentada y de los resultados expuestos, se extraen, a juicio de los autores, algunas importantes aportaciones al campo de la medida en diseños multinivel de medidas repetidas. Se trata, por tanto, de cuestiones que se encuentran en la base de los estudios del valor añadido aplicados a la evaluación educativa, objeto de este monográfico.

En primer lugar, se ha puesto de manifiesto la importancia de conocer y considerar las características diferenciales del patrón de correlaciones de las medidas de rasgo frente a las que frecuentemente se registran en las evaluaciones educativas, que operan sobre la base de puntuaciones observadas. Así, se ha descrito, ejemplificado y simulado, cómo la magnitud de la correlación entre dos medidas es mayor cuanto más se alejan las mismas del punto inicial de la serie de medidas repetidas, siempre y cuando se opere en el nivel de las relaciones causales latentes y se asuma la existencia de un único rasgo. Esto es, a medida que pasa el tiempo y se realizan más aplicaciones o medidas del rendimiento, la correlación entre dos medidas recientes cualesquiera

es mayor que entre dos medidas anteriores en el tiempo, a pesar de que exista la misma distancia temporal entre ellas. Y este patrón se reproduce de un modo estable, sólo con muy ligeras variaciones, cuando se definen diferencialmente otros parámetros tales como la covariación de los residuos en el nivel del alumno o la escuela. Quiere decirse que, en este último sentido, basta con partir en la primera aplicación de una correlación positiva aunque poco elevada entre los residuos de nivel dos y tres para que el paso del tiempo multiplique sus efectos.

En segundo lugar, conocido el patrón de correlaciones, se ha indagado en las razones por las que presumiblemente éste no se refleja en los estudios empíricos ocupados de la evaluación longitudinal del rendimiento. Efectivamente, la literatura asume que cabe esperar una correlación de magnitud similar entre medidas que mantienen la misma distancia temporal. Esto es, que, entre las medidas tomadas en la primera y segunda aplicación, la correlación (típicamente no muy alta, entre 0,7 y 0,8) será la misma que entre las tomadas en la cuarta y la quinta aplicación.

Es innegable que las medidas de rendimiento en los estudios empíricos están íntimamente ligadas con los instrumentos utilizados para medirlas. Sin embargo, cabría esperar que, aún asumiendo que no realizamos medidas perfectas del rasgo latente, *rendimiento en matemáticas* por ejemplo, las correlaciones entre dichas medidas repetidas fueran similares a las que presentan las del rasgo latente, aunque más atenuadas, puesto que asumimos error en la estimación. ¿Qué provoca que esto no siempre ocurra? Una respuesta plausible a esta pregunta es la que se deriva de la consideración de lo que aquí se ha denominado *factor de complejidad*.

Es muy probable que, cuando medimos en repetidas ocasiones un mismo rasgo latente, éste varíe en su naturaleza a lo largo del tiempo. Siguiendo con el ejemplo anterior, el rasgo latente *rendimiento en matemáticas* se convierte en más complejo a medida que el currículo escolar mismo gana en complejidad, exige de aptitudes variadas e incluye facetas diferentes a las tomadas como referente en la primera medición. Esto provoca que el modelo de estimación de las puntuaciones de rasgo a partir de las puntuaciones observadas pierda parte de su eficacia. La dimensionalidad del rasgo latente varía y sin embargo seguimos considerando el rendimiento en matemáticas como el mismo y único rasgo. Parece razonable, entonces, que este hecho explique la ausencia de incremento en la magnitud de las correlaciones entre puntuaciones más alejadas del inicio de la serie de medidas, derivada directamente de una estimación del rasgo progresivamente menos precisa.

Una consecuencia clara de lo anterior es la necesidad de profundizar en el estudio de la estructura dimensional de los rasgos latentes y en su reflejo en la dimensionalidad

de las pruebas con el fin de obtener una medida más aproximada a la realidad del rasgo latente y así hacer una correcta y eficaz estimación. En este sentido, un trabajo de este mismo número (Lizasoain y Joaristi, 2009) pone de manifiesto la importancia de construir y usar pruebas que se ajusten a las características del rasgo, tales como su dimensionalidad.

En esta misma línea de este trabajo, también se deriva lo esencial de la equiparación de puntuaciones. Esto es claro al comparar los patrones de correlación en las puntuaciones latentes y observadas. Para estas últimas, la atenuación expresa la estrecha vinculación de la medida con el instrumento utilizado para obtenerla. Así, puesto que en los estudios longitudinales se obtienen puntuaciones en diferentes momentos y con diferentes instrumentos, la evaluación implica necesariamente contar con ítems que sirvan de anclaje entre aplicaciones consecutivas, con el fin de expresar en la misma escala la habilidad del sujeto.

En tercer lugar, cabe destacar la preocupación constante a lo largo de la historia de la psicometría por la fiabilidad clásica, esto es, por los errores de medida. Los resultados de este trabajo reflejan que, asumiendo fiabilidad constante menor que uno, es decir asumiendo que la varianza de los errores de medida permanece constante a lo largo de las cuatro mediciones, las diferencias entre el patrón de correlaciones de las puntuaciones observadas y el de las puntuaciones verdaderas (eliminando el error de medida) son prácticamente inapreciables. Esto quiere decir que lo importante no es tanto eliminar los errores de medida sino que su variabilidad permanezca constante. En definitiva, las diferencias entre unas aplicaciones y otras no dependen tanto del valor del coeficiente de fiabilidad del instrumento, mientras aseguremos su constancia, como de la creciente complejidad del rasgo que se va a medir.

Entre las posibles limitaciones del trabajo, apuntamos aquí el uso de un modelo concreto de generación de las puntuaciones latentes, con unas características específicas: modelo jerárquico lineal de tres niveles y con dos predictores (el momento de aplicación y el tipo de escuela). Esta elección intenta reflejar las tendencias actuales en la estimación del valor añadido pero es evidente que podrían considerarse modelos no lineales o con más predictores (véase por ejemplo, en este mismo número, Castro, Ruiz y López, 2009; Ferrão, 2009). En este sentido, y también en lo que se refiere a las líneas futuras de trabajo sugeridas por este estudio, destaca el interés de evaluar el efecto de la complejidad creciente de la prueba y de los patrones de covariación de los residuos sobre la estimación misma de medidas del valor añadido.

## Referencias bibliográficas

- BRAUN, H. I. (2005). *Using student progress to evaluate teachers: a primer on value-added models*. Princenton, NJ: Educational Testing Service.
- BHASKAR, R. (1978). *A realist theory of science*. Harvester Press: Sussex.
- CASTRO, M., RUÍZ, C. Y LÓPEZ, E. (2009). Forma básica del crecimiento en los modelos de valor añadido: vías para la supresión del efecto de regresión y funciones de crecimiento no lineales. *Revista de Educación*, 348.
- FERRAÕ, E. (2009). Sensibilidad de las especificaciones de los modelos de valor añadido: la medida del estatus socio-económico. *Revista de Educación*, 348.
- HOUSE, E. (1991). Realism in Research. *Educational Researcher*, 20 (6), 2-9.
- IHAKA, R. Y GENTLEMAN, R. (1996). R: a language for data analysis and graphics. *Journal of Computational and Graphical Statistics*, 5, 299-314.
- LIZASOÁIN, L. Y JOARISTI, L. (2009). Análisis de la dimensionalidad en los modelos de valor añadido: estudio de las pruebas de matemáticas empleando técnicas factoriales y métodos no paramétricos basados en TRI. *Revista de Educación*, 348.
- MARTÍNEZ ARIAS, R. (2009). Usos, aplicaciones y problemas de los modelos de valor añadido en educación. *Revista de Educación*, 348.
- MARTÍNEZ ARIAS, R., GAVIRIA, J. L. Y CASTRO, M. (2009). Concepto y evolución de los modelos de valor añadido en educación. *Revista de Educación*, 348.
- MCCAFFREY, D. L., LOCKWOOD, J. R., KORETZ, D. M. Y HAMILTON, L. S. (2003). *Evaluating value-added models for teacher accountability*. Santa Mónica, CA: RAND Corporation.
- RASCH, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Danish Institute for Educational Research
- SANDERS, W. L. (2006, October 16). *Comparisons among various educational Assessment Value-Added Models*. Paper presented at The Power of Two-National Value-Added Conference, Columbus, Ohio.
- SINGER, J. D. Y WILLET, J. B. (2003). *Applied Longitudinal Data Analysis. Modeling Change and Event Occurrence*. Oxford: Oxford University Press.
- WILEY, E. V. (2006). *A practitioner's guide to value added assessment*. Tempe, AZ: Educational Policy Studies Laboratory, Arizona State University.

**Dirección de contacto:** Coral González Barberá. Universidad Complutense de Madrid. Facultad de Educación. Departamento de Métodos de Investigación y Diagnóstico en Educación. Avenida del Rector Royo Villanova s/n. 28040 Madrid. E-mail: cgbarbera@edu.ucm.es



# Usos, aplicaciones y problemas de los modelos de valor añadido en educación

## Uses, applications and problems of educational value-added models

Rosario Martínez Arias

*Universidad Complutense. Facultad de Psicología. Departamento de Metodología de las Ciencias del Comportamiento. Madrid, España*

### Resumen

El uso de modelos complejos de valor añadido (VA) que intentan aislar la contribución de las escuelas y profesores al desarrollo del estudiante es creciente. Este artículo presenta varios aspectos relacionados con el estado actual de los modelos de VA. Los documentos utilizados para la revisión fueron artículos de revistas con revisión por pares, libros y documentos de instituciones interesadas en los modelos de VA. En primer lugar se presentan algunos modelos actualmente en uso operativo ampliamente aceptados. En segundo lugar, se exponen algunos usos de los modelos de VA. Se identifican dos objetivos principales que se pueden beneficiar del uso de los resultados: mejora de la escuela y rendición de cuentas, con referencia a la elección de escuela. Finalmente se revisa la investigación actual. Se presenta una revisión de los problemas relacionados con el uso de los modelos de VA, que llevan a una interpretación prudente de los resultados. Los problemas se centran en tres aspectos: estadísticos, psicométricos y prácticos. Las cuestiones estadísticas son: sensibilidad a los supuestos de los modelos, estructura del modelo, datos perdidos, temporalidad de los efectos, atribución causal e incertidumbre de los estimadores. Las cuestiones psicométricas están relacionadas con la construcción de escalas verticales para medir el progreso. Los problemas prácticos tienen que ver con la temporalización y elección de las medidas de resultados y con algunas críticas frecuentes como la reducción del currículo y la enseñanza

del test. Se concluye que una aplicación cuidadosa de los modelos de VA, con transparencia y con información adecuada a las partes implicadas, promete ser útil para la evaluación de las escuelas, especialmente para el diagnóstico y mejora y como ayuda en la planificación de las reformas educativas.

*Palabras clave:* rendición de cuentas de las escuelas, mejora de la escuela, problemas estadísticos de los modelos de valor añadido (VA), escalamiento vertical, modelos de valor añadido.

### **Abstract**

The use of complex value-added models (VAM) that attempt to isolate the contribution of schools or teachers to student development is increasing. This paper presents an overview of several aspects related to the state of art of VAMs. The documents used for the review were peer-referees articles, books and reports from different institutions interested in VAMs. Firstly, it presents some VAMs currently used operationally and very well received by stakeholders. Secondly, the paper provides an overview of the main uses of VAMs. Two broad objectives are identified that can benefit from the use of results: school improvement and school accountability with references to school choice. Finally, the current active research is revised. It provides an overview of the issues related to the use of VAM and leads to caution in the interpretation of results. The main issues are related to three topics: statistical, psychometric and practical issues. The statistical issues are: sensitivity of estimated effects to model assumptions, model structure, missing data, volatility of estimated effects, causal attribution and uncertainty of estimates. The main psychometric issues are related to the construction of vertical scales to measure educational progress. The practical issues are related to the timing of measures, choice of measures of results and some frequent critics such as the narrowing of curriculum and the teaching to the test.

We conclude that a diligent application of VAM, with transparency and adequate information to stakeholders, holds considerable promise for school assessment, especially for diagnostic and school improvement, and as aid in the planning of educational reforms.

*Key Words:* school accountability, school improvement, statistical issues of value-added models (VAMs), vertical scaling, value-added models.

## Introducción

Los modelos de valor añadido (VA en adelante) han ganado en popularidad tanto en la comunidad de investigadores, como entre los políticos y administradores escolares. Sus defensores creen que esta nueva forma de plantear la evaluación permite comparar la efectividad de las escuelas, aunque atiendan a poblaciones diversas de estudiantes (Drury y Doran, 2004; McCaffrey, Lockwood, Koretz, y Hamilton, 2003) y guiar los cambios educativos necesarios, tanto en el nivel de la escuela, como de las reformas políticas. Los estimadores numéricos que proporcionan permiten aislar mejor que otros métodos las contribuciones de las escuelas al aprendizaje de los estudiantes, separándolas de otros factores. En Martínez Arias, Gaviria y Castro (2008, en este volumen) se presentan el origen y los principales modelos estadísticos, con un importante cuerpo de investigación que los avala. También existen algunos sistemas de evaluación mediante modelos de VA implantados en la práctica desde hace varios años, que ponen de relieve la factibilidad y utilidad de su aplicación.

No obstante, algunas revisiones (Bock, Wolfe y Fisher, 1996; Kuppermintz, 2003; Mccafrey, et al., 2003), aunque reconocen el importante avance que suponen, sugieren cautelas frente a una aplicación poco crítica, especialmente si los resultados implican consecuencias para profesores y escuelas (Braun, 2005b). Es decir, los modelos de VA, aunque útiles, no son la panacea que permite atribuir de forma inequívoca los resultados del aprendizaje a las escuelas o profesores. Ningún modelo estadístico, ni ningún método de análisis, por sofisticado que sea, puede compensar completamente la falta de aleatorización de los estudios de observación característicos de la investigación sobre los efectos de las escuelas (Braun, 2005a). No obstante, cuando los resultados se interpretan de forma adecuada, como medidas descriptivas, pueden ser de gran utilidad para la evaluación.

Hay algunos modelos que llevan varios años implantados de los que se pueden extraer lecciones sobre los usos potenciales y que mencionaremos a continuación. Podríamos decir que del análisis de estos modelos y de diferentes estudios, se presenta una breve revisión sobre sus usos potenciales. Por otra parte, su aplicación supone el uso de una compleja metodología estadística, enredadas cuestiones de interpretación y una serie de supuestos clave que deben tenerse en cuenta para su aplicación con la «debida diligencia» (Braun, 2005a). En el artículo también se presentan los principales problemas que son objeto de investigaciones actuales sobre el uso estos modelos que deben considerarse para una adecuada aplicación. Los problemas se han agrupado en tres bloques: estadísticos, medida de los resultados mediante tests y prácticos.

Algunos modelos de VA ya han sido aplicados con éxito desde hace años para la monitorización de los rendimientos educativos y para la evaluación de escuelas y profesores. En general han sido valorados positivamente y han tenido una buena aceptación por las partes interesadas (Rose y Gallup, 2007).

Entre los modelos con mayor implantación en la práctica deben mencionarse los siguientes:

- Tennessee Value Added Assessment System (TVAAS, actualmente EVAAS) es el primero en utilizar esta denominación y también en ser utilizado por un estado entero. Ha sido objeto de más estudios y análisis (Amrein-Bearsley, 2008; Ballou, 2002; Bock, Wolfe y Fisher, 1995; Braun, 2005a,b; Braun y Wainer, 2007; Kuppermintz, 2003; McCaffrey, Lockwood, Koretz y Hamilton, 2003). Fue desarrollado por el profesor William Sanders y sus colaboradores (Sanders y Horn, 1994; Sanders, Saxton y Horn, 1997). En la actualidad la aplicación del sistema en los estados que lo requieren se lleva a cabo desde la empresa SAS (que desarrolla el software estadístico), bajo la dirección del profesor Sanders y se denomina *Education Value-Added Assessment System* (EVAAS). Permite obtener perfiles de las escuelas, estableciendo los resultados por quintiles de rendimiento. El método EVAAS proporciona además la denominada «*metodología de proyección*» que proporciona estimadores del nivel de rendimiento del estudiante individual en algún punto futuro, bajo el supuesto de que tendrá una experiencia de escolarización promedio. Wright, Sanders y Rivers (2006) presentan un estudio con dicha metodología. El modelo estadístico es un modelo de los clasificados como de efectos mixtos multivariante y complejo (Martínez Arias et al. 2008), el modelo estratificado. Permite el tratamiento de los datos perdidos, utilizando para cada estudiante todos los datos disponibles. El modelo es muy eficiente en el sentido de que hace uso de toda la información para una cohorte de estudiantes en un período de cinco años. La utilización de tantas medidas y de varias materias minimiza los efectos del error de medida y de la fugacidad o temporalidad de los resultados. No utiliza ajustes de otras covariantes. Esta cuestión ha sido sometida a análisis empíricos y no se ha encontrado como universalmente válida (McCaffrey et al., 2003). No obstante, Ballou, Sanders y Wright (2004) mostraron la posibilidad de inclusión de estos ajustes en el modelo.
- Dallas Value Added Assessment System (Modelo DVAAS): El modelo comenzó en 1984. Parte de una clara filosofía de la equidad, que exige tener en cuenta

para la evaluación las diversas características de los estudiantes y del contexto de la escuela. Desde sus orígenes cuenta con la *Accountability Task Force*, un comité en el que están integradas todas las partes interesadas. Una buena descripción del modelo de evaluación puede encontrarse en Webster y Mendro (1997) y Webster (2005). El modelo sigue longitudinalmente a los estudiantes y utiliza un análisis de datos en dos estadios para los cálculos. En el primer estadio se ajustan las puntuaciones de los dos cursos comparados por variables de los estudiantes, utilizando modelos de regresión de mínimos cuadrados ordinarios. Y los residuos o parte no explicada de las puntuaciones son la base de los cálculos de VA de los profesores y de la escuela seguidos en el segundo estadio, utilizando un modelo lineal jerárquico que controla variables de la escuela.

- Valor añadido de las escuelas públicas de Chicago: El modelo denominado de «productividad» de las escuelas públicas de Chicago evalúa los resultados en competencia lectora y matemáticas de los estudiantes del consorcio entre los cursos 2º y 8º. Se utilizan datos longitudinales de todos los estudiantes para evaluar las ganancias de los estudiantes a lo largo de su permanencia en el sistema educativo. Tiene en cuenta el nivel inicial de los estudiantes y las tendencias de rendimiento. Clasifica las escuelas por medio de un «perfil de productividad». El modelo está descrito en Bryk, Thum, Easton y Luppescu (1998) y es analizado como un modelo multinivel con tres niveles: el desarrollo del estudiante, el estudiante y la escuela. El modelo jerárquico lineal con tres niveles, aunque permite el tratamiento de sujetos con datos perdidos, requiere un anidamiento completo de los estudiantes en la misma escuela para todos los años evaluados, lo que provoca datos perdidos especialmente en las escuelas con alta movilidad. Este hecho les llevó a la introducción de un modelo de clasificación cruzada en los últimos años, lo que permite incorporar la movilidad de los estudiantes. Una descripción del modelo puede encontrarse en Ponisciak y Bryk (2005).
- El modelo de Valor Añadido Contextualizado en Inglaterra: En este volumen Ray, Evans y Mc Cormack (2008) describen la evolución de la experiencia en Inglaterra hasta la actualidad.
- Algunas experiencias españolas: En España, con escasa tradición de evaluación externa de los resultados de los estudiantes, se han desarrollado algunos intentos de evaluación longitudinal con modelos de VA. La experiencia pionera fue la llevada a cabo desde 1997 por el Instituto IDEA a través del Equipo REDES para la evaluación de la Educación Secundaria Obligatoria. Esta experiencia

estuvo limitada a un número reducido de centros voluntarios, en su mayor parte de la Comunidad de Madrid. Una descripción de la experiencia y los principales resultados puede encontrarse en algunos trabajos de miembros del equipo REDES (Marchesi, Martín, Martínez Arias, Tiana y Moreno, 2002; Marchesi, Martínez Arias y Martín, 2004). Una experiencia mucho más ambiciosa y rigurosa es la desarrollada por el equipo investigador de la Facultad de Educación de la Universidad Complutense, dirigido por el profesor Gaviria en la Comunidad de Madrid, con un gran número de centros seleccionados de forma aleatoria. No entramos aquí en su descripción por describirse la experiencia en varios artículos de este volumen (Castro, Ruíz de Miguel y López, 2008; Gaviria, Biencinto y Navarro, 2008; Lizasoán y Juaristi, 2008).

## **Principales usos de los modelos de VA**

Las experiencias realizadas hasta el momento permiten hablar de diferentes usos de los resultados de los modelos de VA. Básicamente pueden agruparse en dos grandes bloques: la rendición de cuentas y la mejora y desarrollo de las escuelas. Los distintos modelos de VA pueden servir para los dos objetivos, pero dado que las connotaciones y consideraciones vinculadas a cada uno de ellos son muy diferentes, conviene establecer claramente los objetivos antes de la introducción de un modelo. Al final del apartado se presenta una breve guía de consideraciones prácticas para un uso fructífero de los modelos.

### **Rendición de cuentas (RC)**

En los últimos años, la adopción de sistemas de rendición de cuentas es frecuente en muchos países, dentro de la tendencia general a la evaluación de resultados en el sector público. El objetivo suele ser la comparación del uso de recursos, resultados y productividad de las diversas instituciones que reciben financiación pública. Los diversos sistemas utilizados pueden tener diferentes objetivos y grados de ambición, que van desde simplemente proporcionar información al gobierno sobre los resultados, servir para identificar buenas prácticas en las instituciones, proporcionar información al público y que los usuarios de los servicios puedan elegir sobre bases informadas,

hasta la aplicación de incentivos directos vinculados a recompensas y sanciones a las instituciones y sus empleados. Sea cual sea la finalidad, conviene tener en cuenta que la rendición de cuentas en el sistema educativo no es un fin en sí misma, sino un medio de lograr las metas educativas. Aunque la rendición de cuentas de una forma u otra es muy antigua, en la actualidad se ha cambiado el foco de atención. En otras épocas se ponía el acento en medidas de recursos y procesos tales como: el control del número de días de instrucción, el tamaño de la clase, las credenciales del profesorado, etc., mientras que los sistemas actuales lo ponen en las medidas de resultados (Hanushek y Raymond, 2004). Un sistema de rendición de cuentas basado en resultados necesita disponer de indicadores simples y equitativos que permitan diferenciar entre escuelas. Los resultados pueden medirse de diferentes formas, pero los modelos de VA utilizan como medida las puntuaciones en tests estandarizados, que juegan un papel dominante debido a su eficiencia de coste-objetividad.

En Martínez Arias et al. (2008) se presentan algunas formas de rendición de cuentas basadas en resultados, especialmente los derivados de la NCLB en Estados Unidos. Los modelos VA pueden usarse para identificar casos extremos, es decir, escuelas que obtienen resultados significativamente mejores o peores de lo esperado. Estas escuelas pueden investigarse posteriormente en relación con sus prácticas docentes, clima escolar, etc. y la información proporcionada puede ser útil para el sistema.

Algunos modelos de VA también evalúan los efectos del profesorado, no obstante, la mayoría de los autores están de acuerdo en que los efectos derivados de los modelos de VA son un indicador muy imperfecto del efecto del profesorado y que en todo caso serán uno de entre múltiples indicadores de su efectividad (Ballou, 2002; Braun, 2005b; McCaffrey et al., 2003).

También existe consenso en que los modelos deben usarse con extraordinaria cautela cuando sus resultados se utilizan en la toma de decisiones con consecuencias para las escuelas o profesores (Ballou, 2002; Braun, 2005b; Kupermintz, 2003; McCaffrey et al., 2003; Raudenbush, 2004a). Los instrumentos y resultados no son lo suficientemente precisos ni fiables para este propósito. Este uso además puede generar fuertes recelos en los profesores y desanimarles a usar la valiosa información que pueden proporcionar los modelos de VA para la mejora de la instrucción (Yeagley, 2007).

Relacionado con la rendición de cuentas está el tema de la publicación de los resultados para facilitar la *elección de escuela*, aunque no suponga incentivos en términos de recompensas o sanciones. Este uso supone proporcionar información a las familias y al público, del rendimiento de las diferentes escuelas para ayudarles en la decisión de la elección.

Los partidarios de la publicidad de los datos y de la elección de escuelas consideran que estas prácticas son un incentivo para proporcionar mejores servicios por parte de las escuelas y elevar el rendimiento de los estudiantes (Glenn y de Groof, 2005; Hoxby, 2000; 2003), entre otras cosas por acoplarse mejor a sus necesidades. Fuchs y Wossman (2007), analizando datos procedentes de evaluaciones internacionales en el nivel de país encontraron que diversas formas de rendición de cuentas, junto con la autonomía y la elección de escuela se relacionaba con niveles más altos de rendimiento de los estudiantes.

En cualquier caso, independientemente de las ventajas y desventajas de la elección, es muy importante cuando se presentan públicamente los resultados de las escuelas que estén en términos de VA contextualizado.

## **Los resultados del VA para la mejora y desarrollo de la escuela**

Se han documentado diversas finalidades para las que las escuelas pueden aplicar los datos de las evaluaciones para la mejora de sus resultados (Supovitz & Klein, 2003). Los datos deben dar respuesta a diversas preguntas, que pueden ejemplificarse en las siguientes: ¿cómo lo estamos haciendo?, ¿proporcionamos una buena educación a los estudiantes? y ¿estamos proporcionando una buena educación a todos los estudiantes? Para poder responder se necesita disponer de información que permita hacer un análisis detallado, comparando el desarrollo de los alumnos en el tiempo, para diferentes asignaturas y grupos de alumnos. Los datos pueden usarse por las escuelas como parte de sus procesos de *auto-evaluación* para diagnosticar las fuerzas y debilidades del centro (Saunders, 2000). En este sentido los resultados de la aplicación del VA pueden ayudar a plantear preguntas sobre las ganancias de los alumnos y a estimular las discusiones informadas entre el personal de los centros sobre la forma de organizar e impartir la enseñanza. Para que estos procesos funcionen bien deben tratarse como innovaciones educativas, proporcionando la preparación adecuada, apoyo y formación para administradores escolares, inspectores, equipos directivos y profesorado.

Dentro de los usos para la mejora de las escuelas es importante destacar el apoyo que los resultados del VA pueden proporcionar a los inspectores mejorando la eficacia de este servicio. Los inspectores, al disponer de esta valiosa información sobre escuelas y estudiantes, pueden guiarles mejor, poniendo el acento en cuestiones clave de la escuela.

## Recomendaciones para el uso adecuado de los modelos de VA

La experiencia ha puesto de relieve usos de los modelos de VA que permiten elevar los rendimientos de los estudiantes. No obstante, para que la experiencia sea fructífera, conviene tener en cuenta una serie de recomendaciones que brevemente se presentan a continuación:

- Clara formulación de los objetivos de la implantación del modelo de VA, ya que diferentes objetivos pueden condicionar el modelo a desarrollar.
- Disponer de datos recogidos sobre los alumnos individuales, con muestras amplias y representativas. Estos datos deben recogerse mediante tests estandarizados externos (estatales o regionales para permitir una adecuada comparación).
- Es conveniente recoger datos de resultados que reflejen todos los niveles de rendimiento del alumnado, no limitándose como en algunos casos de la aplicación de la NCLB a los resultados en torno al nivel de «competente».
- Es imprescindible disponer de varias medidas del rendimiento, y si es posible también del rendimiento en el momento de ingreso en la escuela. En general, como se justifica más adelante, la mayoría de los autores recomiendan la utilización de al menos tres medidas. Como es obvio, es preciso disponer de un identificador único de cada estudiante que se mantiene durante su permanencia en el sistema y que permita vincularlo a la escuela (y al profesor, si el objetivo es evaluar profesorado).
- Aunque como se verá más adelante no existe un consenso generalizado al respecto, es importante disponer de información contextual de los alumnos y de la escuela.
- Utilizar procedimientos de análisis estadístico multinivel, ya que son los únicos que permiten tener en cuenta las dependencias intra-escuela. Estos deben ser preferiblemente del tipo que se ha definido como multivariante (Martínez Arias et al., 2008), ya que permiten analizar los cambios a lo largo del desarrollo y son más eficientes para el tratamiento de valores perdidos.
- Disponer de adecuadas bases de datos y de procedimientos de gestión y control de calidad de las mismas.
- Es imprescindible proporcionar una formación adecuada sobre la interpretación de los datos y sus posibles usos a los equipos directivos de los centros, profesorado y otras partes interesadas.

- Finalmente, debe desarrollarse un sistema de reportado de los resultados claro y transparente. Se recomienda representación tabular, gráfica y verbal, indicando claramente que diferencias son estadísticamente significativas, para no atribuir significados indebidos a las ordenaciones de las escuelas, cuando se establecen. Suele ser conveniente reportar varios tipos de resultados: brutos o no ajustados, ajustados por características del alumnado, como el rendimiento anterior y resultados ajustados por factores contextuales del alumno y de la escuela. También es muy importante en los informes desagregar los resultados por grupos de alumnos, para analizar específicamente sus dificultades y sus ganancias de aprendizaje.

## Las cuestiones estadísticas no resueltas

La aplicación de los modelos de VA requiere de una compleja metodología estadística con fuertes supuestos. En la actualidad muchas cuestiones estadísticas están siendo objeto de investigación y de discusiones entre los científicos. En este apartado se presenta el estado del arte sobre las principales cuestiones objeto de investigación y debate: selección del modelo, la fugacidad o temporalidad de las puntuaciones, ajuste de variables extraescolares, tratamiento de los casos perdidos, la incertidumbre de las estimaciones, cumplimiento de los supuestos de los modelos y las atribuciones causales.

## Las diferencias derivadas del uso de diferentes modelos

Los modelos de VA difieren considerablemente en términos de su complejidad, demandas de datos y facilidad para la comunicación de los resultados. Esta pluralidad de modelos implica que una de las primeras decisiones que se deben tomar es sobre el modelo que se utilizará. Dentro de la elección de modelo, la primera decisión será entre modelos de efectos fijos y de efectos aleatorios. Las aplicaciones derivadas de la aproximación de la función de producción de la economía utilizan sobre todo modelos de efectos fijos (Lockwood y McCaffrey, 2007). En la línea derivada de la evaluación de la efectividad de las escuelas, la mayor parte de los autores optan por la aproximación de efectos mixtos, que tratan los efectos de la escuela como aleatorios

(Ballou, Sanders y Wright, 2004; Sanders, Saxton y Horn, 1997; Goldstein, 2003; McCaffrey, Lockwood, Koretz, Louis y Hamilton, 2004). La investigación ha puesto de relieve que los efectos fijos son muy sensibles a los errores de muestreo, puesto que las escuelas y sobre todo los profesores suelen tener un número reducido de estudiantes lo que lleva a grandes errores muestrales y a la inestabilidad de los estimadores de años sucesivos. Parece más aconsejable la utilización de modelos mixtos de efectos aleatorios, que utilizan estimadores empíricos de Bayes. Estos estimadores tienen la propiedad de ser BLUP (*Best Linear Unbiased Prediction*) y contraen (*shrink*) los estimadores de los efectos de escuelas y profesores hacia la media global según el error o fiabilidad del estimador, consiguiendo una mayor estabilidad de las estimaciones de diferentes años. El problema es que la estimación es más compleja y puede ser más difícil explicarlos a personas sin formación técnica.

Dentro de la aproximación de los modelos mixtos, también existen diversas opciones que difieren en complejidad. La investigación sobre diferencias de resultados derivadas del uso de distintos modelos es todavía escasa y no es concluyente (Ponisciak y Bryk, 2005; Sanders, 2006; Tewke et al., 2004; Wright, 2004; Wright, Sanders y Rivers, 2006).

Cuando la serie de datos es limitada (puntuaciones de dos años) surge la duda de si tratarlos mediante aproximaciones univariantes o modelizar el vector de puntuaciones; parece que la segunda aproximación es más flexible, aunque implique una mayor complejidad en los cálculos.

## Las variables utilizadas en el ajuste

Los modelos de VA permiten ajustar las puntuaciones eliminando los efectos de otros factores extraescolares, que se confunden con los efectos de los profesores y de las escuelas.

Existe un consenso generalizado en que se deben utilizar los niveles de entrada o iniciales de los estudiantes en los ajustes de los modelos de VA, ya que el rendimiento anterior es el mejor predictor del rendimiento futuro (Gray, Jesson, Goldstein, Hedges y Rasbash, 1995; Sammons, Thomas y Mortimore, 1997). En lo que no existe acuerdo es en si se deben ajustar los efectos por otras variables contextuales y en este último caso, qué variables se utilizarán. La mayor parte de los investigadores abogan por el ajuste de dichas variables, aunque el modelo de mayor tradición (TVAAS/EVAAS) lo considera innecesario porque al utilizar múltiples puntuaciones y de varios años, el estudiante sirve como su propio control (Ballou et al., 2004). Parece que la mayor parte de los modelos estadísticos obtienen resultados bastante parecidos

en lo que se refiere a la ordenación de las escuelas (Schafer, Yen y Rahman, 2000; Webster, 2005), pero lo que parece establecer diferencias es la inclusión o no de variables de ajuste y qué variables (Tewke et al., 2004).

No obstante, la postura mayoritaria es la de incluir variables de ajuste, especialmente cuando se establece alguna forma de rendición de cuentas o de difusión de los resultados, ya que la equidad será cuestionable si no se tienen en cuenta características contextuales de los estudiantes y de las escuelas (McCaffrey et al., 2003; McCaffrey et al., 2004).

Partiendo de la necesidad de realizar ajustes, no existe acuerdo sobre qué variables son las importantes, y aún en el caso de algunas consideradas como tales, puede faltar la información adecuada en las bases de datos escolares. La medida de estatus socioeconómico más frecuente de los estudiantes es el disfrute de beca de comedor (*free lunch*), (Ballou et al, 2004; Braun, 2005b; Goldstein, Burgess, y McConnell, 2007; McCaffrey et al., 2004; Sammons et al., 1997), que es una medida bastante imperfecta del estatus, pero frecuentemente es la única disponible. Otras variables presentes en los registros escolares como género, inmigración, etnia, barreras lingüísticas o dificultades de aprendizaje son importantes para el ajuste. La inclusión de dichas variables es importante para analizar el progreso de diferentes subgrupos.

En cuanto a la inclusión de variables contextuales de la escuela, los partidarios de los ajustes lo consideran conveniente. Estas pueden ser agregados de variables de los estudiantes como puntuaciones medias en algunos tests u otras variables definidas en el nivel de la escuela. Entre las más habituales se encuentran: porcentajes de diversos grupos étnicos de interés, estatus socioeconómico de la comunidad, habitat, porcentaje de estudiantes con beca de comedor (utilizado como indicador de pobreza), titularidad de los centros, porcentaje de estudiantes con necesidades educativas especiales, porcentaje de profesores con alta cualificación, niveles de asistencia, etc.

Hay pocos estudios en los que se comparen modelos con diferentes variables de ajuste sobre los mismos datos, pueden consultarse los trabajos de Choi, Goldschmidt y Yamashiro (2006) y de Stevens (2005).

## **La fugacidad o temporalidad de las puntuaciones ganancia**

Cuando los efectos de la escuela se calculan anualmente sobre la base de dos puntuaciones suele haber importantes fluctuaciones en los resultados, lo que ha llevado a advertir sobre el problema de la fugacidad de las ganancias y de las estimaciones

de VA (Kane y Steiger, 2002; Linn y Haug, 2002; Rock, 2007; Way, 2006). El problema se ha explicado con frecuencia por los problemas de la baja fiabilidad de las puntuaciones diferencia (Ballou, 2002; Linn y Haug, 2002) que aumentan los errores de medida. El error de medida y la baja fiabilidad también puede provocar efectos de regresión a la media.

El problema suele paliarse con los modelos multinivel multivariantes con varias mediciones o promediando varias materias (Drury y Doran, 2003; Raudenbush, 2004b; Singer y Willett, 2003), con lo que se aumenta la fiabilidad.

Otros factores responsables de la fluctuación ligados a las medidas de resultados son los cambios en el instrumento de evaluación y el marco (Lockwood y McCaffrey, 2007) y la presencia de efectos techo y suelo en los instrumentos por tener ítems de diferentes cursos o niveles (Rock, 2007).

Los cambios en las ganancias estimadas también se ven afectados por cambios en las variables contextuales y de los estudiantes, así como por el tamaño de las escuelas y cambios en la muestra de escuelas participantes, dado que los resultados siempre se expresan con relación a la muestra. Uno de los aspectos más estudiados ha sido el del tamaño de las escuelas y la mayor variabilidad de las escuelas pequeñas. Este problema se resuelve mejor con el uso de modelos multinivel de efectos aleatorios, ya que su procedimiento de estimación utilizado contraen los estimadores de las escuelas pequeñas, menos fiables, hacia los valores próximos a la media, teniendo en cuenta explícitamente la incertidumbre.

## **El tratamiento de los datos perdidos**

Los modelos de VA requieren muchos datos de los estudiantes. Todos se basan en puntuaciones de los estudiantes en una o más materias durante dos años o más y aquellos que realizan ajustes por características de los estudiantes, necesitan además datos en estas variables. En la práctica, suele haber muchos registros incompletos, lo que produce problemas de datos ausentes en el análisis. Los modelos univariantes requieren la eliminación de todos los casos con datos incompletos; los multivariantes que analizan el vector completo de puntuaciones permiten la inclusión de sujetos con datos perdidos. También se pueden aplicar técnicas de imputación de los valores perdidos antes de los análisis. Sea cual sea la aproximación seguida, la presencia de datos perdidos puede producir importantes sesgos en las estimaciones de los efectos de la escuela.

En la literatura se habla de datos perdidos completamente al azar (MCAR) o al azar (MAR), que en los modelos de VA reflejarían la independencia entre el patrón de los casos perdidos y sus puntuaciones en los tests (Little y Rubin, 2002). Cuando se eliminan los sujetos con datos incompletos, las estimaciones estarán sesgadas a menos que los datos perdidos sigan un patrón MCAR. En los modelos que utilizan todos los casos como TVAAS/EVAAS, parece que no se introducen sesgos si los datos son MAR, lo mismo que cuando se utilizan técnicas de imputación. El problema es que no está claro si los patrones MCAR y MAR se mantienen en los sistemas escolares en los que se aplican los modelos de VA.

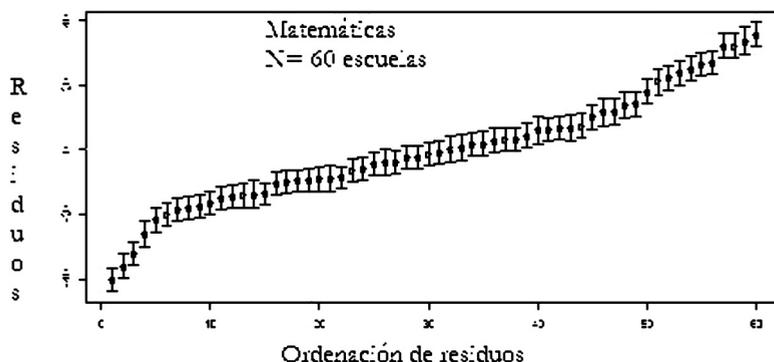
Los investigadores están divididos en cuanto a los efectos de los datos ausentes en las estimaciones de VA, aunque parece que el impacto puede ser importante (Marchesi y Martínez Arias, 2002; Kuppermintz, 2003; McCaffrey et al., 2003; Marchesi et al., 2004; Braun, 2005a; Zvoch y Stevens, 2005a).

## Las incertidumbres derivadas del error muestral

Los modelos de VA producen estimadores de los efectos de la escuela en términos del residuo medio, con valores positivos y negativos y media de 0. Estos efectos deben ir acompañados de su varianza y error típico de estimación para una adecuada interpretación. La varianza estimada de un efecto es una medida del grado de incertidumbre que afecta al estimador y que puede llevar a estimaciones imprecisas (McCaffrey et al., 2003). Esta varianza está determinada en gran medida por el modelo de VA utilizado y por la cantidad de datos disponibles de la escuela o del profesor. Si las varianzas son pequeñas en relación a la amplitud de los efectos, las escuelas «atípicas» pueden identificarse fácilmente y es probable que haya más escuelas estadísticamente diferentes de la escuela promedio. No obstante, existe mucha incertidumbre con las escuelas próximas a los valores medios, lo que hace cuestionable el uso de los modelos de VA para la ordenación de las escuelas. Se recomienda representar los resultados de las escuelas ordenados en un gráfico con bandas de error (Goldstein y Healy, 1995) como el representado en la Figura I, para recoger la incertidumbre de los estimadores y no sobreinterpretar las diferencias.

En la figura, las bandas de error reflejan los intervalos de confianza en torno al efecto y solamente se puede hablar de diferencias estadísticamente significativas entre dos escuelas cuando no se solapan sus bandas. La reducción del error típico es muy importante para poder discriminar entre escuelas y se facilita con conjuntos más

FIGURA I. Residuos (valor añadido) de las escuelas con bandas de confianza



amplios de puntuaciones de los tests y con modelos que hacen un uso más eficiente de los datos. Ballou (2005) encontró bastante imprecisión en la estimación de los efectos del profesor.

### Las atribuciones causales de los efectos de las escuelas

La meta de los modelos de VA es establecer inferencias sobre los efectos de las escuelas o de los profesores (Raudenbush y Willms, 1995, revisado en Martínez Arias et al., 2008). Los modelos introducen una compleja maquinaria estadística para separar dichos efectos de los innumerables factores de confundido que afectan a los resultados y ganancias. Una de las críticas frecuentes es que la palabra efecto puede tener una connotación causal, en el sentido de que son las escuelas o los profesores la causa de los resultados encontrados en sus estudiantes (McCaffrey et al., 2003; Rubin, Stuart y Zanutto, 2004). Aunque la inferencia causal es la meta última y en la que estarían interesados los políticos y administradores escolares, es evidente que no se puede extraer este tipo de interpretaciones a partir de estudios observacionales como los utilizados en los modelos de VA (Kuppermintz, 2003; Raudenbush, 2004b; Rubin et al., 2004). El reto de la estimación de los efectos es la comparación entre un resultado observado y un resultado potencial. El resultado observado es la puntuación o la ganancia de un estudiante que ha asistido a una escuela  $j$ . El resultado potencial es el que se habría observado si el mismo estudiante hubiese sido asignado a una condición alternativa  $j'$ . Evidentemente, el estudiante solamente es observado en la condición

j y el problema de los modelos de VA es encontrar el resultado potencial para dicho estudiante. Este resultado ausente suele reemplazarse por el promedio de otras escuelas a las que hubiera podido asistir el estudiante. Este resultado podría ser un adecuado «contrafactual» si la asignación de estudiantes y profesores a las escuelas se realizase de forma aleatoria, de modo que todas estuviesen sirviendo a poblaciones similares, lo que es un supuesto falso. Las inferencias causales solamente pueden extraerse de experimentos aleatorizados con adecuados grupos de control, ya que la aleatorización y grandes muestras reduce la probabilidad de las influencias de las múltiples variables de confundido. El problema de los modelos de VA es que no disponen de diseño experimental que permita evaluar efectos causales. Interpretar un estimador estadístico derivado de un estudio observacional sin aleatorización como una medida directa de un efecto causal es cuestionable (Rosenbaum, 2002; Raudenbush, 2004b; Rubin et al. 2004; Stuart, 2007), ya que puede haber numerosas influencias fuera del control de la escuela. Los modelos de VA intentan captar las virtudes de los experimentos aleatorizados cuando no se han realizado, ya que los datos de la escuela provienen de un estudio de observación y no de un experimento. Son un intento de compensar la ausencia de aleatorización, ajustando por características de los estudiantes y contextuales, estableciendo supuestos de que las covariantes que influyen en los resultados captan adecuadamente las diferencias entre condiciones. El investigador identifica un residuo que permanece en la escuela después de ajustar otras variables que pueden explicar los resultados. Para alejar la tentación de interpretación causal, algunos autores prefieren llamarlos «efectos residuales de la escuela/profesor» (Fitz-Gibbon, 1997; Schatz, Von Secker y Alman, 2005). Por muy complejo que sea el modelo estadístico, no puede compensar la falta de aleatorización. La investigación en ciencias sociales está repleta de ambigüedades en la interpretación, reflejadas claramente en la paradoja de Lord (Lord, 1967). Por este motivo existe un acuerdo generalizado en huir de las atribuciones causales, poniendo el acento en los aspectos descriptivos proporcionados por los estimadores (Raudenbush, 2004b; Rubin et al., 2004; Thum, 2006), aunque estas descripciones son muy útiles para hacer proposiciones sobre el progreso de los estudiantes y de las escuelas.

## **El cumplimiento de los supuestos de los modelos**

La mayor parte de los procedimientos estadísticos utilizados en los modelos de VA son procedimientos paramétricos que requieren el cumplimiento de numerosos supuestos que condicionan la seguridad de los estimadores. En la realidad los supuestos

nunca se cumplen completamente y ningún modelo es perfectamente adecuado. La cuestión es determinar la dirección y la magnitud de los sesgos frente a su violación. Los supuestos pueden referirse a la naturaleza de los datos, la estructura del modelo o a ambos. Por lo que se refiere a la estructura de los datos, al igual que en los modelos de regresión suele asumirse la normalidad y la homocedasticidad. También es importante la ausencia de efectos techo y suelo en los datos, ya que su presencia puede conducir a distribuciones asimétricas.

El hecho de trabajar con puntuaciones de tests que presentan errores de medida, puede afectar también a los estimadores. En los modelos en los que se ajustan covariantes como predictores, sean puntuaciones de los tests de años anteriores o de otras variables, un supuesto es que las variables están medidas sin error y el uso de predictores con error afecta a los estimadores de la regresión, pudiendo tener un efecto sustancial en la estimación de los efectos de la escuela. En el caso de las puntuaciones de los tests, el efecto puede reducirse tomando puntuaciones de varios años (Ladd y Walsh, 2002; Lockwood y McCaffrey, 2007; Sanders, 2004). La presencia de errores de medida también afecta a otras variables utilizadas en el ajuste que en muchas ocasiones proceden de autoinformes de los estudiantes y están bastante afectadas por el error. Los errores de medida no uniformes también pueden llevar a la heterocedasticidad de los resultados, contribuyendo a estimaciones sesgadas (McCaffrey et al., 2003).

Por otra parte, en los modelos multivariantes, se requiere más investigación sobre los efectos de los diferentes tratamientos de las estructuras de covarianza (Raudenbush y Bryk, 2002; Singer y Willett, 2003).

En resumen, es necesario realizar más investigación sobre la robustez de los estimadores en presencia de incumplimiento de supuestos, que algunos autores califican como «heroicos» (Rubin et al., 2004).

## **Cuestiones relacionadas con las medidas de resultados en los modelos de VA**

Los modelos de VA tienen como principal soporte las puntuaciones en tests de rendimiento educativo y sus inferencias por lo tanto requieren de tests con adecuadas propiedades psicométricas. La utilidad de su información depende de la calidad de los tests empleados. Desgraciadamente, aunque hay tests mejores que otros, podemos afirmar que «no existe una medida perfecta del rendimiento», proposición con la que están de acuerdo incluso los más ardientes defensores de los tests.

Uno de los requisitos de gran parte de los modelos es el de disponer de las denominadas *escalas verticales* y, tal como señalan McCaffrey et al. (2003), los cambios en los procedimientos de construcción, los métodos usados en la vinculación de los diferentes tests que las componen o el peso relativo dado a sus componentes, pueden influir en las inferencias sobre los efectos de las escuelas o los profesores. En Martínez Arias et al. (2008), este volumen se presentan algunos modelos en los que no se requieren tests verticalmente escalados, como algunos basados en estándares de rendimiento, que utilizan puntuaciones ordinales y que se podrían tratar con los modelos multinivel ordinales Fielding, Yang, Goldstein (2003). No obstante, la evaluación en estos casos, aunque más sencilla, también está basada en puntuaciones de los tests (Lissitz y Huyhn, 2003; Huyhn y Schneider, 2005) y deben gozar de adecuadas propiedades psicométricas. También es posible transformar los datos de los diferentes tests a otros tipos de escalas normativas, como las puntuaciones equivalentes en curva normal (Gong, Perie y Dunn, 2006; McCall, Kingsbury y Olson, 2004; Smith y Yen, 2006), pero las escalas verticales parecen las más adecuadas con los modelos de VA (Choi, Goldschmidt y Yamashiro, 2006).

En general, los modelos parten del supuesto de que las puntuaciones son *bastante buenas*, pero sin entrar en demasiados detalles sobre lo que esta afirmación significa y considerando el tema de las puntuaciones poco problemático.

Lissitz y Huyhn (2003) definen el escalamiento como el proceso mediante el que las puntuaciones directas obtenidas en un test se transforman en un nuevo conjunto de números, con algunos atributos particulares como media y desviación típica. Los mismos autores definen la *escala vertical* como una escala única unidimensional que resume el rendimiento de los estudiantes de diferentes cursos o niveles. Estas escalas permiten establecer comparaciones del rendimiento de los estudiantes a lo largo de diversos cursos para estimar su progreso (Patz, 2007).

La construcción de escalas verticales requiere utilizar procedimientos conocidos en la literatura estadística como de equiparación de formas distintas de tests, con el objetivo de que sus puntuaciones sean comparables. Una breve descripción de los principales métodos y diseños puede encontrarse en Martínez Arias, Hernández Lloreda y Hernández Lloreda (2006) y en Kolen (2006). Un tratamiento completo en Kolen y Brennan (2004) y en Dorans, Pommerich y Holland (2007). Básicamente, hay dos situaciones bien diferenciadas en las que se necesita la conversión de un conjunto de tests a una escala común o equiparada. La equiparación horizontal tiene lugar cuando se usan formas múltiples, de dificultades similares, para evaluar un contenido semejante dentro del mismo curso o edad. La equiparación vertical es denominada

*vinculación vertical (vertical linking)* por no cumplir estrictamente con los requisitos de la equiparación (Holland, 2007) y se utiliza para vincular puntuaciones de tests que evalúan el mismo contenido general, pero con niveles de dificultad diferentes, normalmente destinados a distintos cursos. Las puntuaciones de escala resultantes se denominan escalas verticales. El escalamiento vertical es el proceso de vincular diferentes niveles de un instrumento que miden el mismo constructo, en una escala común de puntuaciones (Kolen, 2006; Kolen y Brennan, 2004; Lissitz y Huyhn, 2003; Patz, 2007; Petersen, Kolen y Hoover, 1989). La validez de las inferencias del VA sobre las ganancias depende en gran medida de la adecuación de las escalas, que generalmente disminuye cuando aumenta la distancia entre cursos o niveles.

Hay numerosas cuestiones prácticas no resueltas relativas a estas escalas y muy pocas recomendaciones claras sobre su construcción. De hecho, los *Standards for Educational and Psychological Tests* (AERA, APA & National Council on Measurement in Education, 1999), las mencionan solamente de pasada. En su construcción deben asumirse diferentes supuestos y tomarse diversas decisiones que pueden llevar a diferentes resultados en los modelos de VA que las utilizan (Harris, 2007; Patz, 2007; Patz y Yao, 2007a; Patz y Yao, 2007b; Tong y Kolen, 2007; Yen, 2007). Su aplicación, cada vez más frecuente en los modelos de VA, ha llevado a una reflexión crítica sobre sus supuestos. Para un análisis más detallado de estos problemas se recomienda la lectura de las referencias citadas, especialmente Harris.

## **Alternativas a las escalas verticales convencionales**

La alternativa más prometedora es utilizar en su construcción modelos psicométricos más complejos, que permitan considerar constructos multidimensionales que reflejen los cambios curriculares y de desarrollo. La opción son los modelos multidimensionales de la teoría de la respuesta al ítem (*Multidimensional Item Response Theory*, MIRT). En este sentido ya existen resultados de algunos de estos modelos (Martineau et al., 2007; Reckase y Li, 2007; Reckase y Martineau, 2004; Patz y Yao, 2007a, 2007b; Roberts y Ma, 2006). No obstante, en las aplicaciones prácticas puede haber algunas dificultades como la identificación de la dimensionalidad adecuada y los procedimientos requeridos para la vinculación. Otra dificultad añadida es la falta de software comercial.

Otra opción es considerar el rendimiento como medida ordinal, utilizando los niveles definidos por los estándares, como en algunos de los modelos presentados en

Martínez Arias et al. (2008). En este caso el problema es el de la comparabilidad entre los niveles de rendimiento de los distintos cursos y para garantizarla se han propuesto los *estándares verticalmente moderados* (Cizek, 2005; Huyhn y Schneider, 2005; Lissitz y Huyhn, 2003) o *verticalmente articulados* (Ferrara, Phillips, Williams, Leinwand, Mahoney y Ahadi, 2007). Básicamente, como señalan Huyhn y Schneider (2005) sus dos elementos básicos son: un conjunto de definiciones políticas de los niveles de rendimiento que se usan para todos los cursos y una línea de tendencia consistente entre cursos que se impone a los porcentajes de estudiantes en las diferentes categorías. Razones de espacio impiden entrar aquí en la descripción de los procedimientos, pero una buena lectura es número monográfico de la revista *Applied Measurement in Education* (2005, vol.v18, núm. 1 ) coordinado por Cizek.

## Cuestiones prácticas ligadas a la aplicación y uso de los tests

### La administración de los tests

En cuanto a la aplicación de los tests de rendimiento hay dos cuestiones importantes: el número de medidas de resultados y las fechas de recogida de datos. En relación con la primera cuestión, un sistema de evaluación de VA implica al menos dos mediciones de los resultados, normalmente de dos cursos consecutivos. Los modelos con dos evaluaciones consecutivas analizan la diferencia como ganancia o utilizan la primera como predictora de la segunda. Aunque estos modelos son bastante comunes, no permiten todo el potencial de los estudios longitudinales, ya que no permiten examinar la forma de la función de desarrollo. Como Rogosa (1995) advierte «dos medidas son mejor que una, pero no mucho mejor» (p. 744). Además en estos casos, pueden producirse algunos de los problemas ya tratados en el apartado de los problemas estadísticos, por lo que se recomienda utilizar medidas múltiples. Estas permiten un tratamiento multivariante y multinivel con el que se pueden estimar en el primer nivel los parámetros del cambio y desarrollo y obtener una visión más estable de las ganancias de la escuela Drury y Doran, 2003; Raudenbush, 2004a, 2004b; Raudenbush y Bryk, 2002).

La segunda cuestión se refiere al momento de la recogida de datos. En muchas de las aplicaciones de los modelos de VA se recogen datos una vez al año, normalmente

en primavera. Algunos autores cuestionan estas fechas de aplicación, ya que el intervalo entre aplicaciones incluye partes de dos años académicos con la interrupción del verano, cuyos efectos pueden estar relacionados con el estatus socioeconómico de los estudiantes. McCaffrey et al. (2003) consideran que esto no supone una amenaza a la validez de las inferencias. Otra alternativa considerada a veces es la realización de dos recogidas de datos en el curso, en otoño y primavera. Esta opción, además de más costosa, es desaconsejada por algunos autores (Linn, 2000) puesto que puede introducir sesgos derivados de la selección de los estudiantes, errores de conversión de las escalas y efectos de la práctica. En los modelos basados solamente en dos puntuaciones, es importante la elección de la primera ocasión de medida, ya que el punto inicial puede afectar a la interpretación de los resultados.

### **Los incentivos «perversos» derivados del uso de los tests**

Las evaluaciones mediante tests que tienen como objetivo la rendición de cuentas o la publicidad de los resultados son objeto de críticas frecuentes por parte de algunos investigadores. Aunque el repertorio de críticas es extenso, la mayor parte se centran en dos grandes grupos: la reducción del currículo y la inflación de las puntuaciones de los tests.

El primer tipo de críticas suele referirse a las materias evaluadas y a los tests utilizados en la evaluación. Por lo que se refiere a las materias evaluadas, la crítica se basa en que muchos modelos de VA utilizados en la rendición de cuentas solamente evalúan un número reducido de materias, siendo las más frecuentes las Matemáticas y la Lengua (las requeridas por la NCLB en Estados Unidos). Esta restricción de contenidos puede enfatizar las materias objeto de evaluación y descuidando las restantes. En realidad, esta crítica es aplicable a cualquier sistema de evaluación mediante tests y no solamente a los modelos de VA, ya que, teóricamente pueden construirse tests de cualquier materia, si bien es cierto que las diferencias curriculares entre cursos en Ciencias Naturales y Ciencias Sociales hacen más difícil la construcción de escalas verticales.

Otra crítica frecuente en relación con la anterior es la que insiste en que la evaluación para la rendición de cuentas lleva a los docentes a utilizar estrategias conocidas como «enseñanza para el test» (Kohn, 2000; O'Day, 2002). La crítica adquiere mayor virulencia cuando los tests utilizan formatos de elección múltiple de los que se dice que no captan las destrezas de solución de problemas y procesos superiores. Haciéndose eco de esta crítica, cada vez más evaluaciones incluyen cuestiones de respuesta construida y de ensayo, aunque algunos problemas relacionados con su calidad

técnica y sobre todo con los costos, impiden utilizarlos en las evaluaciones sistemáticas con periodicidad frecuente (Wang et al., 2006). Jacob (2002, 2005) no encontró evidencias de reducción del currículo en las escuelas de Chicago.

Debido a las prácticas docentes vinculadas a los contenidos de los tests, se puede producir la *inflación de las puntuaciones* (Koretz, 2005; Koretz y Hamilton, 2006; Linn, 2006), que lleva a incrementos en las puntuaciones de los tests que no representan auténticas ganancias de aprendizaje y que no se reflejan en otras evaluaciones.

En el fondo de todas estas discusiones se encuentra la denominada Ley de Campbell (Campbell, 1975, p. 35) «cuanto más se usa un indicador social cuantitativo para la toma de decisiones, más sometido estará a presiones de corrupción y habrá mayor tendencia a distorsionar y viciar el proceso social que intenta monitorizar». La generalización de la ley de Campbell está tan extendida que Madaus y Clark (2001) la comparan al principio de incertidumbre de Heisenberg en las Ciencias Sociales. Nichols y Berliner (2005) hacen una excelente revisión de evidencias de la Ley de Campbell en diversas áreas de las ciencias sociales, entre otras en la evaluación educativa, mostrando en este ámbito ejemplos de efectos como el estrechamiento del currículo, la enseñanza del test, la subjetividad en el establecimiento de estándares y la exclusión de las sesiones de tests de algunos grupos de estudiantes.

Otra crítica frecuente, se refiere al carácter de evaluación normativa de los modelos de VA (los resultados o residuos se basan en la comparación con otras escuelas de la muestra). En opinión de algunos investigadores, este tipo de evaluación puede rebajar los niveles de rendimiento de los estudiantes con relación a los que se obtendrían por medio de una evaluación basada en estándares. Se demanda poder responder a la pregunta *¿cuánta ganancia es suficiente?* En realidad, la respuesta a esta pregunta no es incompatible con los modelos de VA que siguen durante varios años el aprendizaje de los estudiantes y permiten proyectar cuáles serán sus puntuaciones. Algunos de los modelos citados en Martínez Arias et al. (2008) intentan compatibilizar ambas aproximaciones.

## Conclusiones y recomendaciones para la investigación futura

Mediante una compleja metodología estadística, los modelos de VA parecen proporcionar un procedimiento adecuado para la evaluación de la eficacia de las escuelas, ya que separa hasta cierto punto la contribución de las escuelas y profesores de otros

factores extraescolares. Con la utilización de medidas longitudinales del desarrollo de los estudiantes proporcionan fundamentos más defendibles que los tradicionales métodos transversales basados en el estatus de la escuela en un momento en el tiempo. Estas propiedades de los VA los han llevado a ser considerados como una promesa para realizar evaluaciones de escuelas de forma más justa y segura. Por ello han despertado un fuerte interés tanto entre los investigadores, como entre los políticos y administradores escolares.

Existen ya varios modelos implantados desde hace años cuya experiencia permite reflexionar sobre sus potencialidades de uso, fuerzas y debilidades. Aunque tienen algunas limitaciones técnicas y prácticas, en general han tenido una buena aceptación entre las diferentes partes implicadas en la educación.

Los principales usos de los modelos pueden resumirse en los relacionados con la rendición de cuentas y con el diagnóstico y mejora de las escuelas.

En cuanto al primero de los usos, la aplicación de los resultados del VA no parece adecuada por el momento cuando tiene consecuencias para las escuelas o los profesores. Hay numerosas cuestiones no resueltas sobre la atribución clara de los efectos mostrados en los residuos de los modelos a las escuelas. La complejidad de las escuelas nunca puede ser totalmente captada con los modelos de VA como para atribuirles efectos causales sobre el aprendizaje de sus estudiantes. Los modelos de ganancias simples son aún más problemáticos para esta atribución que los más complejos.

Si la rendición de cuentas consiste simplemente en la publicación de sus resultados para información al público y ayuda a los padres en la elección de escuela, los datos basados en VA son más justos que los resultados brutos no ajustados. No obstante, hay que tener cautela en la difusión de la información, especialmente cuando se establecen ordenaciones, puesto que hay pocas diferencias estadísticamente significativas entre la mayor parte de las escuelas, debido a la incertidumbre o error muestral.

Los modelos de VA pueden resultar muy útiles en los procesos de diagnóstico y mejora de las escuelas. Permiten identificar escuelas que necesitan asistencia, ayuda y formación del profesorado y son un componente esencial en los procesos de autoevaluación de los centros para el diagnóstico de sus fuerzas y debilidades. Combinados con otras metodologías de naturaleza más cualitativa como observaciones, entrevistas, portafolios, etc., ayudan a la identificación de buenas prácticas que se pueden potenciar en las reformas educativas.

Desde el punto de vista técnico, los modelos de VA son un área de investigación activa de la que son esperables nuevos desarrollos metodológicos que permitan mejorar tanto los modelos como los instrumentos de medida utilizados en la evaluación

de resultados. Hay algunos aspectos concretos que necesitan más investigación: estudios empíricos de comparación de diferentes modelos, influencia del uso de diferentes características contextuales, efectos de la no aleatoriedad de los datos perdidos, efectos de la violación de los supuestos estadísticos y sensibilidad de los modelos a los diferentes métodos de construcción de tests y a los efectos de sus propiedades psicométricas.

Se requieren además investigaciones empíricas sobre otros aspectos apenas investigados hasta el momento y sobre los que los modelos de VA son frecuente objeto de críticas: validación externa de los efectos, contrastándolos con otras medidas de eficacia de escuelas y profesores y sobre los supuestos «incentivos perversos» a los que según algunos lleva su aplicación.

Otros aspectos muy importantes que hay que mejorar para el uso eficaz de los datos son los que tienen que ver con la comunicación e interpretación de los resultados por parte de audiencias que no son expertas en estadística. La interpretación de los residuos derivados de complejos modelos estadísticos como efectos de las escuelas puede ser difícil y poco transparente. Es imprescindible llegar a informes que sean de fácil comprensión por las escuelas y profesores, insistiendo en la transparencia de los resultados y en la utilidad para los procesos de decisión de la escuela. Para ello, la implantación de los modelos de VA deberá ir acompañada de planes de formación para el profesorado y otros implicados.

En resumen, a pesar de algunas de las limitaciones presentadas en este trabajo, muchas de las cuales están siendo objeto de investigación para su mejora, los modelos de VA, utilizados de forma adecuada y prudente, sin la pretensión de hacer inferencias para las que por el momento no están capacitados, representan un valioso instrumento al servicio de las escuelas.

## Referencias bibliográficas

- AERA, APA & NACIONAL COUNCIL ON MEASUREMENT IN EDUCATION (1999). *Standards for Educational and Psychological Tests*. Washington, DC: American Psychological Association.
- AMREIN-BEARDSLEY, A. (2008). Methodological concerns about the education value-added assessment system. *Educational researcher*, 37, 65-75.
- BALLOU, D. (2002). Sizing up test scores. *Education Next*, 2, 10-15.

- (2005). Value-added assessment: Lessons from Tennessee. En R. LISSITZ (ed.), *Value-Added models in education: Theory and Applications* (pp. 272-297). Mapple Grove, MN: JAM Press
- BALLOU, D., SANDERS, W. & WRIGHT, P. (2004). Controlling for student background in value-added assessment of teachers. *Journal of Educational and Behavioral Statistics*, 29, 37-65.
- BOCK, R. D., WOLFE, R. & FISHER, T. H. (1996). *A review and analysis of the Tennessee value-added assessment system. Summary and recommendations*. Nashville, TN: Tennessee Office of Education Accountability.
- BRAUN, H. (2005a). Value-Added Modeling: What Does Due Diligence Require? En R. LISSITZ (ed.), *Value-Added models in education: Theory and Applications* (19-39). Mapple Grove, MN: JAM Press.
- (2005b). *Using Student Progress to Evaluate Teachers: A Primer on Value-Added Models*. Princeton, NJ: Educational Testing Service.
- BRAUN, H. & WAINER, H. (2007). Value-added modelling. En S. SINHARAY & C. RAO (eds.), *Handbook of statistics, 26: Psychometrics* (pp. 867-891). Amsterdam: North Holland.
- BRYK, S.A., THUM, M.Y., EASTON, Q.J. & LUPPESCU, S. (1998). *Academic productivity of Chicago public elementary schools*. Chicago, IL: Consortium on Chicago School Research.
- CAMILLI, G. (1988). Scale shrinkage and the estimation of latent distribution parameters. *Journal of Educational Statistics*, 13, 227-241.
- CAMPBELL, D.T. (1975). Assessing the impact of planned social change. En G. LYONS (ed.), *Social research and public policies: The darmouth/OECD Conference* (pp. 3-45). Hanover, NH: Dartmouth College. The Public Affairs Center.
- CASTRO, M., RUÍZ DE MIGUEL, C. Y LÓPEZ, E. (2008). Forma básica del crecimiento en los modelos de valor añadido: vías para la supresión del efecto de regresión y funciones de crecimiento no lineales. *Revista de Educación*, 348.
- CHOI, K., GOLDSCHMIDT, P. & YAMASHIRO, K. (2006). *Exploring models of school performance: From theory to practice* (CSE Rep: No. 673). Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing.
- CIZEK, J. (ed.). (2005). Adapting testing technology to serve accountability aims: The Case of vertically-moderated standard setting. A Special Issue of *Applied Measurement in Education*, 18, 1-9.
- CIZEK, G. J. & BUNCH, M. B. (2007). *Standard setting: a guide to establishing and evaluating performance standards on tests*. Thousand Oaks, CA: Sage.
- COHEN, J., JOHNSON, E. & ANGELES, J. (2000). *Variance estimation when sampling in two dimensions via the jackknife with application to the national assessment of educational progress*. Washington, DC: American Institute for Research.

- DORAN, H. C. (2004). *Value-Added Models and Adequate Yearly Progress: Combining Growth and Adequacy in a Standards-based Environment*. Paper presented at the Annual CCSSO Large-Scale Assessment Conference, Boston, Massachusetts.
- DORAN, H. C. & COHEN, J. (2005). The confounding effect of linking bias on gains estimated from value-added models. En R. LISSITZ (ed.), *Value-added models in education: Theory and applications* (pp. 80-110). Maple Grove, MN: JAM Press.
- DORAN, H. C. & FLEISCHMAN, S. (2005). Research matters/Challenges of value-added assessment. *Educational Leadership*, 63, 85-87.
- DORAN, H. C. & IZUMI, L.T. (2004). *Putting Education to the Test: A Value-Added Model for California*. San Francisco, CA: Pacific Research Institute.
- DORAN, H. C. & JIANG, T. (2006). The impact of linking error in longitudinal análisis: an empirical demonstration. En R. LISSITZ (ed.), *Longitudinal and value added models of student performance* (pp. 210-229). Maple Grove, MN: JAM Press.
- DORAN, H. C. & LOCKWOOD, J. R. (2006). Fitting value-added models in R. *Journal of Educational and Behavioral Statistics*, 31, 205-230.
- DORANS, N. J., POMMERICH, M. & HOLLAND, P. W. (eds.). (2007). *Linking and aligning scores and scales*. New York: Springer.
- DRURY, D. & DORAN, H. (2003). The Value of Value-Added Analysis. *NSBA Policy Research Brief*, 3, 1-4.
- FERRARA, S., PHILLIPS, G. W., WILLIAMS, P. L., LEINWAND, S., MAHONEY, S. & AHADI, S. (2007). Vertically articulated performance standards: An exploratory study of inferences about achievement and growth. En R. LISSITZ (ed.), *Assessing and modeling cognitive development in school* (pp. 31-63). Maple Grove, MN: JAM Press.
- FIELDING, A., YANG, M. & GOLDSTEIN, H. (2003). Multilevel ordinal models for examination grades. *Statistical Modelling*, 3, 127-153.
- FITZ-GIBBON, C. T. (1997). *The value-added national project: Final report: feasibility studies for a national system of value added indicators*. London: School Curriculum and Assessment Authority.
- FUCHS, T. & WOSSMANN, L. (2007). What accounts for international differences in student performance? A re-examination using PISA data. *Empirical Economics*, 32, 433-464.
- GAVIRIA, J. L., BIENCINTO, M. C. Y NAVARRO, E. (2008). Invarianza de la estructura de covarianzas de las medidas de rendimiento académicos en estudios longitudinales en la transición de la educación primaria a secundaria. *Revista de Educación*, 348.
- GLENN, C. & DE GROOF, J. (2005). *Balancing Freedom, Autonomy and Accountability in Education*. Nijmegen NL: Wolf Legal Publishers.

- GOLDSCHMIDT, P. & CHOI, K. (2007). *The practical benefits of growth models for accountability and the limitations under NCLB. Policy Brief 9*. Los Angeles: National Center for Research on Evaluation, Standards, and Student Testing.
- GOLDSCHMIDT, P., CHOI, K. & MARTINEZ, F. (2004). *Using Hierarchical Growth Models To Monitor School Performance Over Time: Comparing NCE to Scale Score Results* (CSE Report 618). Los Angeles, CA: Center for Research on Evaluation Standards and Student Testing.
- GOLDSTEIN, H. & HEALY M. J. R. (1995). The Graphical Presentation of a Collection of Means. *Journal of the Royal Statistical Society*, 581, 175-177.
- GOLDSTEIN, H. (2003). *Multilevel models*. London: Arnold.
- GOLDSTEIN, H., BURGESS, S. & MCCONNELL, B. (2007). Modelling the effect of pupil mobility on school differences in educational achievement. *Journal of The Royal Statistical Society*, A, 170, 941-954.
- GRAY, J., JESSON, D., GOLDSTEIN, H., HEDGER, K. & RASBASH, J. (1995). A multilevel analysis of school improvement: changes in schools' performance over time. *School Effectiveness and School Improvement*, 10, 97-114.
- GULLIKSEN, H. (1950). *Test theory*. New York: Wiley.
- HAERTEL, E. H. (2004). *The behavior of linking items in test equating* (CSE Report No 630). Los Angeles, CA: Center for Research on Evaluation Standards and Student Testing.
- HANSON, B. A. & BEGUIN, A. A. (2002). Obtaining a common scale for IRT item parameters using separate versus concurrent estimation in the common item non-equivalent groups equating design. *Applied Psychological Measurement*, 21, 3-24.
- HARRIS, D. J. (2007). Practical issues in vertical scaling. En N. J. DORANS, M. POMERICH, & P. W. HOLLAND (eds.), *Linking and aligning scores and scales* (pp. 233-251). New York: Springer.
- HILL, R., SCOTT, M., DE PASCALE, C., DUNN, J. & SIMPSON, M. A. (2006). *Using value tables to explicitly value student growth*. En R. LISSITZ (ed.), *Longitudinal and value added models of student performance* (pp. 255-290). Maple Grove, MN: JAM Press.
- HOLLAND, P. W. (2007). *A framework and history for score linking*. En N. J. DORANS, M. POMERICH & P. W. HOLLAND (eds.), *Linking and aligning scores and scales* (pp. 5-30). New York: Springer.
- HOXBY, C. (2000). Does competition among public schools benefit students and taxpayers. *American Economic Review*, 90, 1.209-1.238.
- (2003). School choice and school competition: Evidence from the United States. *Swedish Economic Policy Review*, 10, 9-65.

- HUYHN, H. & SCHNEIDER, C. (2005). Vertically moderated standards: Background, assumptions, and practice. *Applied Measurement in Education*, 18, 99-114.
- JACOB, B. (2002). *Accountability, Incentives and Behavior: The Impact of High-Stakes Testing in the Chicago Public Schools*. Cambridge, MA.: NBER Working Paper No. 8968.
- KANE, T. J. & STAIGER, D. O. (2002). *Volatility in school test scores: Implications for test-based accountability systems*. En D. Ravith (ed.), *Brooking papers on education policy* (pp. 235-260). Washington, DC: Brookings Institution.
- KIM, S. & COHEN, A. S. (1998). A comparison of linking and concurrent calibration under item response theory. *Applied Psychological Measurement*, 22, 131-143.
- KOHN, A. (2000). *The case against standardized testing: Raising the scores, ruining the schools*. Portsmouth, NH: Heineman.
- KOLEN, M. J. (2006). Scaling and norming. En R. L. BRENNAN (ed.). *Educational Measurement* (4ª ed.). (pp. 155-186). Westport, CT: American Council on Education.
- KOLEN, M. & BRENNAN, R. (2004). *Test equating, scaling, and linking: Methods and practices* (2ª ed.). New York: Springer-Verlag.
- KORETZ, D. M. (2005). Alignment, high stakes and the inflation of test scores. En J. L. HERMAN & E. H. HAERTEL (eds.), *Uses and misuses of data in accountability testing. Yearbook of the National Society for the Study of Education*. vol. 104 Part 2 (pp. 99-118). Malden, MA: Blackwell.
- KORETZ, D. M. & HAMILTON, L. S. (2006). Testing for accountability in K-12. En R. L. BRENNAN (ed.), *Educational Measurement* (4ª ed.). (pp. 531-578). Westport, CT: Praeger.
- KUPERMINTZ, H. (2003). Teacher Effects and Teacher Effectiveness: A Validity Investigation of the Tennessee Value Added Assessment System. *Educational Evaluation and Policy Analysis*, 25(3).
- LADD, H. F. & WALSH, R. P. (2002). Implementing value-added measures of school effectiveness: Getting the incentives right. *Economics of Education Review*, 21, 1-17.
- LINN, R. L. (2000). Assessments and Accountability. *Educational Researcher*, 29, 4-14.
- (2006). Validity of inferences from test-based educational accountability systems. *Journal of Personnel Evaluation in Education*, 19, 5-15.
- LINN, R. L. & HAUG, C. (2002). Stability of school building accountability scores and gains. *Educational Evaluation and Policy Analysis*, 24, 29-36.
- LISSITZ, R. (ed.). (2005). *Value added models in education: Theory and applications*. Maple Grove, MN: JAM Press.
- (2006). *Longitudinal and Value added models of student performance*. Maple Grove, MN: JAM Press.

- LISSITZ, R. W. & HUYNH, H. (2003). Vertical equating for state assessments: issues and solutions in determination of adequate yearly progress and school accountability. *Practical Assessment, Research, and Evaluation*, 8, 10.
- LITTLE R. J. A. & RUBIN D. B. (2002). *Statistical Analysis with Missing Data* (2ª ed.) New York: Wiley.
- LIZASOAIN, L. Y JUARISTI, L. (2008). Análisis de la dimensionalidad en los modelos de valor añadido: estudio de pruebas de matemáticas empleando técnicas factoriales y métodos no paramétricos basados en TRI. *Revista de Educación*, 348.
- LOCKWOOD, J. R. & MCCAFFREY, D. F. (2007). Controlling for individual heterogeneity in longitudinal models with applications to student achievement. *Electronic Journal of Statistics*, 1, 223-252.
- MADAUS, G. & CLARKE, M. (2001). The adverse impact of high-stakes testing on minority students: Evidence from one hundred years of test data. En G. ORFIELD & M. L. KORNHABER (eds.), *Raising standards or raising barriers? Inequality and high-stakes testing in public education* (pp. 85-106). New York: Century Foundation Press.
- MARCHESI, A., MARTÍN, E., MARTÍNEZ ARIAS, R., TIANA, A. & MORENO, A. (2002). An evaluation network for educational change. *Studies in Educational Evaluation*, 29, 43-56.
- MARCHESI, A., MARTÍNEZ ARIAS, R. Y MARTÍN, E. (2004). Estudio longitudinal sobre la influencia del nivel sociocultural en el aprendizaje de los alumnos de la ESO. *Infancia y Aprendizaje*, 27, 307-323.
- MARTINEAU, J. A. (2004). *The effects of construct shift on growth and accountability models*. Unpublished Doctoral Dissertation. Michigan State University.
- (2006). Distorting value-added: The use of longitudinal, vertically scaled student achievement data for growth-based value-added accountability. *Journal of Educational and Behavioral Statistics*, 31, 35-62.
- MARTINEAU, J. A., SUBEDI, D. R., WARD, K. H., LI, T., LU, Y., DIAO, Q., PANG, F. H., DRAKE, S., SONG, T., KAO, S. C., ZHENG, Y. & LI, X. (2007). Non-linear trajectories through multidimensional content spaces: An examination of psychometric claims of unidimensionality, linearity, and interval measurement. En R. LISSITZ (ed.), *Assessing and modeling cognitive development in school* (pp. 96-142). Maple Grove, MN: JAM Press.
- MARTÍNEZ ARIAS, R., GAVIRIA, J. L. Y CASTRO, M. (2008). Concepto y evolución de los modelos de valor añadido en educación. *Revista de Educación*, 348.
- MARTÍNEZ ARIAS, R., HERNÁNDEZ LLOREDA, V. Y HERNÁNDEZ LLOREDA, M. J. (2006). *Psicometría*. Madrid: Alianza.
- MCCAFFREY, D. F., LOCKWOOD, J. R., KORETZ, D. M. & HAMILTON, L. S. (2003). *Evaluating Value-Added Models for Teacher Accountability*. Santa Mónica, CA: RAND Corporation.

- MCCAFFREY, D. M., LOCKWOOD, J. R., KORETZ, D., LOUIS, T. A. & HAMILTON, L. (2004). Models for value-added modeling of teacher effects. *Journal of Educational and Behavioral Statistics*, 29, 67-101.
- MCCALL, M. & HAUSER, C. (2007). Item response theory and longitudinal modeling: The real world is less complicated than we fear. En R. LISSITZ (ed.), *Assessing and modeling cognitive development in school* (pp. 143-174). Maple Grove, MN: JAM Press.
- MICHAELIDES, M. P., & HAERTEL, E. H. (2004, May). *Sampling of common items: An unrecognized source of error in test equating* (Technical Report). Los Angeles: Center for the Study of Evaluation y National Center for Research on Evaluation, Standards, and Student Testing.
- O'DAY, J. (2002). Complexity, Accountability, and School Improvement. *Harvard Educational Review*, 72, 293-329.
- PATZ, R. J. (2007). *Vertical scaling in standards-based educational assessment and accountability systems*. Washington, DC: The Council of Chief State School Officers.
- PATZ, R. J. & YAO, L. (2007a). Vertical scaling: Statistical models for measuring growth and achievement. En S. SINHARAY & C. RAO (eds.), *Handbook of statistics, 26: Psychometrics* (pp. 955-975). Amsterdam: North Holland.
- (2007b). Methods and models for vertical scaling. En N. J. DORANS, M. POMMERICH & P.W. HOLLAND (eds.), *Linking and aligning scores and scales* (pp. 253-272). New York: Springer.
- PETERSEN, N. S., KOLEN, M. J. & HOOVER, H. D. (1989). Scaling, norming and equating. En R. L. LINN (ed.), *Educational Measurement*. (3<sup>a</sup> ed.). (221-262). New York: Macmillan.
- POMPLUM, M., OMAR, M. H. & CUSTER, M. (2004). A comparison of Winstep and Bilog-MG for vertical scaling with the Rasch model. *Applied Psychological measurement*, 28, 247-273.
- PONISCIAK, S. M. & BRYK, A. S. (2005). Value-added analysis of the Chicago public schools: An application of hierarchical models. En R. LISSITZ (ed.). *Value-Added models in education: Theory and applications* (pp. 40-79). Mapple Grove, MN: JAM Press.
- RAUDENBUSH, S. (2004a). *Schooling, statistics, and poverty: Can we measure school improvement?* Princeton, NJ: Educational Testing service.
- (2004b). What are value-added models estimating and what does this imply for statistical practice? *Journal of Educational and Behavioral Statistics*, 29, 121-129.
- RAUDENBUSH, S. W. & WILLMS, J. D. (1995). The estimation of school effects. *Journal of Educational and Behavioral Statistics*, 20, 307-335.

- RAY, A., EVANS, H. & McCORMACK, T. (2008). The use of national value added models for school improvement in English schools. *Revista de Educación*, 348.
- RECKASE, M. D. (2004). The real world is more complicated than we would like. *Journal of Educational and Behavioral Statistics*, 29, 117-120.
- RECKASE, M. D. & LI, T. (2007). Estimating gain in achievement when content specifications change: A multidimensional item response theory approach. En R. LISSITZ (ed.), *Assessing and modeling cognitive development in school* (pp. 189-204). Maple Grove, MN: JAM Press.
- RECKASE, M. D. & MARTINEAU, J. (2004). *The vertical scaling of science achievement tests*. Paper commissioned by the Committee on Test Design for K-12 Science Achievement. Washington, DC: National Research Council.
- ROBERTS, J. S. & MA, Q. (2006). *IRT models for the assessment of change across repeated measurements*. En R. LISSITZ (ed.), *Longitudinal and value added models of student performance* (100-129). Maple Grove, MN: JAM Press.
- ROGOSA, D. (1995). Myths about longitudinal research. En J. M. GOTTMAN (ed.), *The analysis of change* (pp. 3-66). Mahwah, NJ: Erlbaum.
- ROSENBAUM, P. R. (2002). *Observational studies*. New York: Springer.
- ROSE, L. C. & GALLUP, A. M. (2007). The 39<sup>th</sup> Annual Phi Delta Kappa/Gallup poll of the public's attitudes toward the public schools. *Phi, Delta, Kappan*, 89, 33-48.
- RUBIN, D. B., STUART, E. A. & ZANUTTO, E. E. (2004). A potential outcomes view of value-added assessment in education. *Journal of Educational and behavioural Statistics*, 29, 103-116.
- SAMMONS, P., THOMAS, S. & MORTIMORE, P. (1997). *Forging links: Effective schools and effective departments*. London: Chapman y Hall.
- SANDERS, W. & HORN, S. (1994). The Tennessee value-added assessment system (TVAAS): Mixed model methodology in educational assessment. *Journal of Personnel Evaluation*, 9, 299-311.
- SANDERS, W. L., SAXTON, A. M. & HORN, S. P. (1997). The Tennessee value-added assessment system: a quantitative, outcomes-based approach to educational assessment. En J. MILLMAN (ed.), *Grading teachers, grading schools: Is student achievement a valid evaluation measure?* (pp. 137-162). Thousand Oaks, CA: Corwin Press.
- SAUNDERS, L. (2000). Understanding schools' use of value-added data. The psychology and sociology of numbers. *Research Papers in Education: Policy and Practice*, 15, 241-258.
- SCHMIDT, W. H., HOUANG, R. T. & MCKNIGHT, C. C. (2005). Value-added research: Right idea but wrong solution? En R. LISSITZ (ed.), *Value-added models in education: Theory and applications* (pp. 145-165). Maple Grove, MN: JAM Press.

- SCHAFFER, W. D., YEN, S. J. & RAHMAN, T. (2000). School effects indices: Stability of one-and-two level formulations. *Journal of Experimental Education*, 68, 239-250.
- SCHATZ, C. J., VONSECKER, C. E. & ALBAN, T. R. (2005). Balancing accountability and improvement: Introducing value-added models to a large school system. En R. LISSITZ (ed.), *Value added models in education: Theory and applications* (pp. 1-18). Maple Grove, MN: JAM Press.
- SMITH, R. L. & YEN, W. M. (2006). Models for evaluating grade-to-grade growth. En R. W. LISSITZ (ed.), *Longitudinal and value added modeling of student performance* (pp. 82-99). Maple Grove, MN: JAM Press.
- STEVENS, J. (2005). The study of school effectiveness as a problem in research design. En R. LISSITZ (ed.), *Value-Added models in education: Theory and applications* (pp. 166-208). Maple Grove, MN: JAM Press.
- STUART, E. A. (2007). Estimating causal effects using school-level datasets. *Educational Researcher*, 36 (4), 187-198.
- SUPOVITZ, J. A. & KLEIN, V. (2003). *Mapping a course for improved student learning*. Philadelphia: Consortium for Policy Research in Education.
- TEKWE, C. D., CARTER, R. L., MA C.-X., ALGINA, J., LUCAS, M., ROTH, J. et al. (2004). An empirical comparison of statistical models for value-added assessment of school performance. *Journal of Educational and Behavioral Statistics*, 29, 11-36.
- THUM, Y. M. (2006). Designing gross productivity indicators: A proposal for connecting accountability goals, data and analysis. En R. LISSITZ (ed.), *Longitudinal and value added models of student performance* (pp. 436-479). Maple Grove, MN: JAM Press.
- TONG, Y. & KOLEN, M. J. (2007). Comparison of methodologies and results in vertical scaling for educational achievement tests. *Applied Measurement in Education*, 20, 227-253.
- WANG, L., BECKETT, G. H. & BROWN, L. (2006). Controversies of standardized assessment in school accountability reform: A critical synthesis of multidisciplinary research evidence. *Applied Measurement in Education*, 19, 305-328.
- WEBSTER, W. J. (2005). The Dallas school level accountability model: The marriage of status and value-added approaches. En R. LISSITZ (ed.), *Value added models in education: Theory and applications* (pp. 233-271). Maple Grove, MN: JAM Press.
- WEBSTER, W. J. & MENDRO, R. L. (1997). The Dallas value-added accountability system. En J. MILLMAN (ed.), *Grading teachers, grading schools: Is student achievement a valid evaluation measure?* (81-99). Thousand Oaks, CA: Corwin Press.
- WRIGHT, S. P. (2004). Advantages of a multivariate longitudinal approach to educational value-added assessment without imputation. Paper presented at the 2004 National Evaluation Institute, July 8-10, Colorado Springs, Colorado.

- WRIGHT, S. P., SANDERS, W. L. & RIVERS, J. C. (2006). Measurement of academic growth of individual students toward variable and meaningful academic standards. En R. LISSITZ (ed.), *Longitudinal and value added models of student performance* (pp. 385-406). Maple Grove, MN: JAM Press.
- YANG, M., GOLDSTEIN, H., RATH, T. & HILL, N. (1999). The use of assessment data for school improvement purposes. *Oxford Review of Education*, 25, 469-483.
- YEN, W. M. (1986). The choice of scale for educational measurement: An IRT perspective. *Journal of Educational Measurement*, 23, 299-325.
- (2007). Vertical scaling and No Child Left Behind. En N. J. DORANS, M. POMMERICH & P. W. HOLLAND (eds.), *Linking and aligning scores and scales* (pp. 273-283). New York: Springer.
- YEN, W. M. & BURKET, G. R. (1997). Comparison of item response theory and Thurstone methods of vertical scaling. *Journal of Educational Measurement*, 34, 293-313.
- YEN, W. M. & FITZPATRICK, A. R. (2006). Item response theory. En R. L. BRENNAN (ed.), *Educational Measurement* (4ª ed.). (pp. 111-154). Westport, CT: Praeger.
- ZVOCH, K. & STEVENS, J. J. (2005). Sample exclusion and student attrition effects in the longitudinal study of middle school mathematics performance. *Educational Assessment*, 10, 105-123.

## Fuentes electrónicas

- GOLDSCHMIDT, P., ROSCHEWSKI, P. CHOI, K., AUTY, W., HEBBLER, S., BLANK, R. & WILLIAMS, A. (2005). *Policymakers' guide to growth models for school accountability: How do accountability models differ?* Washington, DC: Councils of Chief State School Officers. Consultado de: <http://www.ccsso.org/publications/>
- GONG, B., PERIE, M. & DUNN, J. (2006). *Using Student Longitudinal Growth Measures for School Accountability Under No Child Left Behind: And Update to Inform Design Decisions*. Consultado el 9 de enero de 2008, de [http://www.nciea.org/publications/GrowthModelUpdate\\_BGMAPJD07.pdf](http://www.nciea.org/publications/GrowthModelUpdate_BGMAPJD07.pdf)
- LISSITZ, R., DORAN, H., SCHAFER, W. & WILLHOFT, J. (2006). Growth modeling, value-added modeling and linking: An introduction. En R. LISSITZ (ed.), *Longitudinal and Value-Added Models of Student Performance* (pp. 1-46). Maple Grove, Minnesota: JAM Press. Consultado de: <http://www.nwea.org>
- MCCALL, M. S., KINGSBURY, G. G. & OLSON, A. (2004). *Individual growth and school success*. Lake Oswego, OR: Northwest Evaluation Association. Consultado de: <http://www.nwea.org>

- NICHOLS & BERLINER (2005). *The inevitable corruption of indicators and educators through high-stakes testing*. Tempe, Arizona: Arizona State University. Education Policy studies Laboratory. Consultado de: <http://edpolicylab.org>
- SANDERS, W. L. (2006). Comparisons Among Various Educational Assessment Value-Added Models. The Power of Two-National Value-Added Conference, Columbus, OH. Consultado el 20 de mayo de 2007. Consultado de: <http://www.sas.com/govedu/edu/services/vaconferencepaper.pdf>
- SCHAFER, W.D. (2006). Growth scales as an alternative to vertical scales. *Practical Assessment, research y Evaluation*, 11(4). Consultado de: <http://pareonline.net/getvn.asp?v=11yn=4>
- YEAGLEY, R. (2007). Separating Growth from Value Added: Two Academic Models Offer Different Tools for Different Purposes. *The School Administrator*, 64(1). Consultado el 18 de octubre de 2007. Consultado de: <http://www.aasa.org/publications/saarticledetail.cfm?ItemNumber=7941>

**Dirección de contacto:** Rosario Martínez Arias. Universidad Complutense de Madrid. Departamento de Metodología de las Ciencias del Comportamiento. Facultad de Psicología. Campus de Somosaguas, 28223 Pozuelo de Alarcón, Madrid. E-mail: [rmnez.arias@psi.ucm.es](mailto:rmnez.arias@psi.ucm.es)



# **Investigaciones y estudios**



# Los sistemas de reconocimiento y acreditación de los aprendizajes no formales e informales: referencias históricas, funciones socioeducativas y perspectiva teórica

## The recognition and accreditation of non-formal and informal learning: historical references, social functions and theoretical perspective

Óscar Medina Fernández

*Universidad de las Palmas de Gran Canaria. Facultad de Formación del Profesorado. Las Palmas de Gran Canaria, España*

Florentino Sanz Fernández

*Universidad Nacional de Educación a Distancia. Madrid, España.*

### Resumen

Los sistemas de reconocimiento y acreditación de los aprendizajes no formales e informales no solo se están convirtiendo en una práctica habitual en los países desarrollados, sino que están obligando a cambiar buena parte de los postulados y certidumbres que se mantienen actualmente en el campo de la educación. Desde el punto de vista conceptual, hablamos de un juicio evaluador que consiste en reconocer y certificar las competencias que posee una persona, independientemente de cuándo, dónde y cómo se han adquirido.

En la práctica se trata de la actuación de un tribunal que evalúa determinados aprendizajes que equivalen a las competencias de una titulación oficial. Estos aprendizajes se han podido adquirir mediante procesos de educación formales, no formales o informales.

En este trabajo se presenta una parte de los resultados de una investigación sobre *Sistemas de Acreditación*. En primer lugar, se hace un breve recorrido histórico sobre estas prácticas educativas en Estados Unidos, Canadá y Europa, deteniéndonos en la situación actual española.

En segundo lugar, se analizan las funciones y ventajas que estos sistemas de acreditación tienen para los trabajadores, para las empresas y para la sociedad y el sistema educativo en general. Finalmente, se plantean algunos elementos de reflexión que creemos que están en la base del debate teórico sobre estos nuevos sistemas de titulación: los problemas que plantean, los principios y características que se comparten por la mayoría de los profesionales, las perspectivas teóricas desde las que cabe abordarlos y su inevitable relación con los conceptos de educación formal, no formal e informal.

*Palabras clave:* acreditación, sistemas de acreditación, reconocimiento de la competencia, valoración de aprendizajes no formales e informales, aprendizaje a lo largo de la vida, educación permanente, educación de adultos, formación profesional.

### **Abstract**

The recognition and accreditation of learning acquired in informal and non-formal contexts are not only becoming standard practice in developed countries, but they are also imposing changes on a great number of long-held beliefs and attitudes which are still present in the field of education today. From a conceptual point of view, we are referring to evaluating criteria consisting of the recognition and certification of the individual competencies that a person has, regardless of when, where and how they have been acquired.

In practical terms, an expert group made up of professionals from the formal educational system will evaluate certain learning experiences that are equivalent to the competences of an official qualification. These competences could be acquired through formal, non-formal or informal learning.

This paper shows some of the results of an investigation on *Accreditation Systems*. Firstly, we conduct a brief historical overview of these educational practices in USA, Canada and Europe, examining the current situation in Spain more closely. Secondly, we analyse the purposes and advantages that these accreditation systems bring to workers, enterprises, society and, generally, to the education system. And finally, some particular aspects are considered for further reflection, as we consider them to be the basis for a theoretical discussion about these new qualification systems: problems that arise from them, principles and features shared by most professionals, the theoretical framework from which they should be approached, and their inevitable connection with formal, non-formal and informal education.

*Key Words:* accreditation, accreditation systems, recognition of competencies, validation of non-formal and in-formal learning, lifelong learning, permanent education, adult education, vocational training.

## Introducción

El tema que nos ocupa constituye actualmente un fenómeno educativo de actualidad en el que se están depositando no pocos anhelos y expectativas. No en vano, se trata de un asunto analizado en los últimos años en múltiples reuniones, encuentros y congresos, y sobre el que además se publican numerosos artículos y documentos.

Se piensa en los sistemas de reconocimiento y acreditación de la competencia como una respuesta a los nuevos problemas que tiene la educación en la sociedad actual; se entiende que este nuevo quehacer educativo va a permitir abordar con más rigor las desigualdades asociadas a los procesos de formación; desde el punto de vista de la producción y el trabajo el reconocimiento de la experiencia está aportando más transparencia a los mercados laborales; algunos incluso hablan de estos sistemas como un nuevo recurso para la mejora de los sistemas educativos formales y como el gran reto de la educación de las personas adultas para el siglo XXI.

Con todo, no faltan las reservas e incluso los análisis negativos. Se oyen voces que hablan de los sistemas de reconocimiento como una suerte de aventura pedagógica que podría representar una amenaza para la unidad y coherencia del sistema educativo. También están los que rechazan estos sistemas por sus perversos efectos sociales, toda vez que pueden ocasionar la descalificación social de los títulos, o un estribero fácil para titular y colocarse en el mercado laboral. Otras críticas descalifican de antemano la posibilidad de que a través de la experiencia se puedan adquirir competencias suficientes sobre una profesión o un conjunto estructurado, ordenado y jerarquizado de conocimientos y aprendizajes.

Estamos pues ante un tema controvertido, que en unos despierta entusiasmos y en otros desconfianza. Una cuestión sobre la que se está planteando un importante debate prácticamente en todos los países desarrollados y que está desatando diversas y acaloradas disputas entre profesionales y expertos.

Pero ¿de qué hablamos cuando usamos la expresión que da título a este artículo? Podemos responder con un caso real. Antonio tiene 46 años y es jefe de cocina de un hotel. Comenzó a trabajar a los 16 años ya terminada la EGB. Ayudaba en las tareas más elementales de la cocina del hotel. Durante sus 26 años de experiencia laboral siempre se ha interesado por aprender y mejorar en su trabajo, fijándose en cómo actuaban los demás, preguntando lo que no sabía y tratando de hacer siempre el trabajo lo mejor posible, a la vez que no desaprovechaba la ocasión de asistir a cursos ocupacionales o cursos impartidos por la propia empresa sobre cocina y manipulación de alimentos. De este modo su buen hacer y su competencia profesional se han ido incre-

mentando, lo que le ha permitido ir mejorando su categoría laboral hasta que le han nombrado jefe de cocina.

Evidentemente, si a Antonio le han nombrado jefe de cocina será porque tiene las competencias de un cocinero. La dirección de un hotel no toma una decisión como ésta si no ha comprobado suficientemente sus competencias y la capacidad de liderazgo que la cocina de un hotel lleva consigo.

En resumen, estamos ante un profesional de la cocina cuyas competencias y conocimientos no los ha adquirido mediante la formación inicial-formal, sino mediante la experiencia y las enseñanzas no formales. Dicho de otra manera, sabe y domina lo mismo que un titulado, pero no tiene el título.

El caso de Antonio no es extraño ni único. Todos los sectores laborales, y casi todas las empresas tienen muchos casos como el de Antonio en su plantilla, que fundamentalmente han ido aprendiendo a través de la experiencia laboral y de enseñanzas no formales.

Si Antonio es evaluado por un tribunal y comprueba que, en efecto, tiene las competencias, podrá certificar que las posee y proponerle para el título de Técnico en Cocina, del mismo modo que los enseñantes de esta titulación aprueban y titulan a los alumnos que forman y superan los exámenes.

En este caso lo que hace el tribunal es aplicar un sistema de evaluación cuya finalidad es reconocer y acreditar en términos de enseñanza formal los aprendizajes que se han adquirido por vías no formales e informales. En la práctica se trata de un proceso que pasa por varias fases: información, orientación, evaluación, certificación e itinerario formativo. A esto es a lo que se les llama *sistemas de reconocimiento y acreditación de los aprendizajes no formales e informales*, aunque hay otras expresiones para hablar de lo mismo: sistemas de acreditación, validación de lo adquirido, validación de competencias, balance de competencias, o también, nuevos sistemas de titulación, etc.

Usemos una u otra expresión, siguiendo con el caso de Antonio, hablamos de reconocer en términos de enseñanza formal (dado que por esta vía se puede adquirir un título oficial), los aprendizajes y competencias que se han adquirido mediante sistemas no formales (como, por ejemplo, cursos ocupacionales) o informales (como la experiencia y los procesos de autoaprendizaje).

Por tanto, conceptualmente hablando, nos referimos a los *procesos de evaluación en virtud de los cuales se reconoce y se certifica que una persona posee determinadas competencias independientemente de cómo, dónde y cuándo se han adquirido*. Entendemos que ésta es la esencia de lo que estamos hablando: reconocer y

certificar lo que una persona sabe (los aprendizajes adquiridos, las competencias), no importa cómo, ni dónde, ni cuándo. En el fondo, todo esto significa que un título se puede adquirir mediante la formación inicial formal (ésta ha sido la vía tradicional) y también mediante los sistemas de acreditación consistentes en reconocer y dar el mismo valor educativo a los procesos de educación no formales e informales (aquí está la novedad).

Presentamos en este trabajo una parte de los resultados de la investigación sobre «Sistemas de Acreditación», financiada por la Universidad de Las Palmas de Gran Canaria (proyectos de investigación del programa propio de la ULPGC, convocatoria de 28 de abril de 2005). Se trata de un proyecto de investigación que nos ha permitido indagar sobre múltiples aspectos y perfiles de los sistemas de acreditación, entre otros: el marco conceptual, las funciones socioeducativas, su desarrollo en otros países y en España, el análisis de experiencias, la acreditación en el ámbito universitario, el desarrollo experimental de métodos y protocolos, así como sus fundamentos epistemológicos, históricos, psicológicos, sociales y pedagógicos.

No obstante, en este artículo nos limitamos a tres cuestiones que nos parecen relevantes. En primer lugar, se hace un breve recorrido histórico sobre estas prácticas en Estados Unidos, Canadá y Europa, deteniéndonos en la situación actual de este fenómeno en España. En segundo lugar, se analizan las funciones y ventajas que estos sistemas de acreditación tienen para los trabajadores, para las empresas y para la sociedad y el sistema educativo en general. Finalmente, se plantean algunos elementos de reflexión que creemos que están en la base del debate teórico sobre estos nuevos sistemas de titulación: los problemas que plantean, los principios y características que se comparten por la mayoría de los profesionales, las posibles perspectivas teóricas desde las que cabe abordarlos y su relación inevitable con los conceptos de educación formal, no formal e informal.

## Referencias históricas

Aunque en España resulta muy novedoso el impulso que, en las últimas leyes educativas, se está dando a la necesidad de reconocer y validar los aprendizajes no formales e informales; este fenómeno no es, sin embargo, nada nuevo en los países del entorno occidental. En Estados Unidos y Canadá el reconocimiento y la acreditación de los

aprendizajes, independientemente de su modo de adquisición, tienen una tradición de casi un siglo y en algunos países europeos la preocupación por este tema se hace presente desde mediados del siglo pasado con el advenimiento de la denominada sociedad del conocimiento,

## Estados Unidos y Canadá

El primer movimiento norteamericano a favor del reconocimiento de los aprendizajes adquiridos extrainstitucionalmente lo protagonizaron los militares de la Primera Guerra Mundial al reivindicar una equivalencia entre la formación que habían recibido en el ejército y la que se recibía en las instituciones de formación colegial. Esta reivindicación no tuvo respuesta, sin embargo, hasta que los soldados de la Segunda Guerra Mundial, apoyados por amplios colectivos de adultos, forzaron al Estado, a las instituciones de enseñanza, a los servicios comunitarios y a las empresas a pactar una solución sobre los aprendizajes adquiridos en la experiencia. Con los soldados de la guerra del Vietnam, años después, se volvería a repetir una reivindicación similar (Beaudet, 1984, p. 157).

En este país la terminología utilizada es tan variada como la que encontramos en otros lugares: se habla de aprendizajes extraescolares, aprendizajes anteriores, aprendizajes experienciales, aprendizajes extrainstitucionales, aprendizajes de formación extraescolar, etc. Creemos que la denominación no es una cuestión de menor importancia, aunque en todos los casos parece estar claro que el factor común es el que se refiere a los aprendizajes independientemente de las condiciones y modalidades de su adquisición.

Otro de los hitos en la historia de Estados Unidos sobre la acreditación fue el protagonizado por los movimientos feministas, que veían muy sesgadas las normas pactadas en el primer programa de reconocimiento de aprendizajes. Criticaban que sólo se reconocieran aprendizajes adquiridos en oficios masculinos, en perfiles ocupacionales de la vida pública y en puestos de trabajo remunerados, aspectos que no siempre se cumplían en las experiencias adquiridas por las mujeres (Meyer et al., 1987, p. 51).

Hay que esperar a los años setenta, década de las mujeres, declarada por la ONU, para que tengan éxito las primeras reivindicaciones de reconocimiento de aprendizajes específicos adquiridos por las mujeres en EEUU. Ya no se va a tratar de reconocer sólo los aprendizajes adquiridos en actividades que ocupan los hombres, ni sólo en perfiles laborales asalariados y públicos: se incluirán también los aprendizajes adquiridos en el

hogar, en el compromiso de voluntariado y en la propia experiencia de vida. Las mujeres consideraban que en todos estos lugares se adquirirían conocimientos y competencias transferibles a la vida pública y a la vida productiva asalariada.

En Canadá, el reconocimiento de los aprendizajes adquiridos en la experiencia (RAE) también tiene un largo recorrido. Al igual que en Francia, los canadienses usan el familiar término *acquis* referido a los aprendizajes no formales e informales. En este país lo que da lugar a que se comience a hablar sobre el reconocimiento de competencias es el informe de la Comisión de Estudios sobre la Formación de Adultos (CEFA), presidida por Michèle Jean (1982) y conocido popularmente como «Informe de la Comisión Jean». Se trata de un amplio estudio sobre la educación de adultos en Canadá desde la perspectiva de la educación permanente. El capítulo segundo de la cuarta parte se dedicó íntegramente a analizar lo que significaba el reconocimiento y sus consecuencias para los individuos que quieren insertarse en el mercado laboral o continuar estudios en el sistema de enseñanza (Meyer et al., 1978, p. 115).

El debate sobre el reconocimiento de aprendizajes adquiridos se generaliza en todo el Canadá a través de múltiples conferencias y foros. En 1986 se celebra la primera conferencia anual sobre el RAE y en 1995 el primer foro nacional al que asisten más de 500 personas. Cuando, en 1994, se crea la Asociación Canadiense para el Reconocimiento de lo Adquirido (CAPLA) será esta asociación la que coordine la celebración sistemática de conferencias y foros nacionales sobre el reconocimiento de los aprendizajes hasta el año 2006. Desde la Comisión Jean se ha avanzado mucho en Canadá en el tema de reconocimiento de los aprendizajes, ha habido experiencias originales y se han consolidado criterios, dispositivos y procedimientos.

## **Países europeos**

En Europa, el reconocimiento de aprendizajes no formales e informales tiene menos de un siglo de historia, aunque con resultados muy desiguales según los países. Los primeros ensayos por validar y acreditar competencias profesionales adquiridas al margen de los sistemas formativos surgen, en algunos países europeos, en los entornos de la Primera Guerra Mundial cuando los psicólogos alemanes intentaron evaluar las aptitudes de los ciudadanos para conducir vehículos motorizados en tiempos de guerra con el objetivo de reclutar personal capacitado para ayudar al ejército alemán. Después de la guerra extendieron la aplicación de estas pruebas psicotécnicas a otros ámbitos profesionales como, por ejemplo, el ferroviario, necesitado de personal especialmente competente.

Otras experiencias mucho más cercanas a las que conocemos en la actualidad sobre validación de aprendizajes no formales son las realizadas en los años treinta (1934) por las «Escuelas Especiales» francesas con obreros no cualificados que, desempeñando tareas propias de ingenieros, terminaron reivindicando y consiguiendo el reconocimiento y la acreditación de sus competencias. En los años setenta, y siempre dentro del marco legislativo de la formación profesional francesa, se comienza a hablar del derecho social al reconocimiento y de la obligación nacional a satisfacerlo (Ley francesa del 16 de julio de 1971) y mucho más tarde aún se hablará también de reconocer y validar no solamente los aprendizajes adquiridos en la experiencia de trabajo sino en la experiencia de participación social. La primera vez que se aporta en Francia este nuevo matiz de validar también los aprendizajes adquiridos en la vida social es cuando la comisión permanente del Consejo Nacional de Formación Profesional adapta para Francia la definición del *portfolio* canadiense en el año 1985. El *portfolio* es definido como un conjunto de pruebas que describen las competencias adquiridas por un individuo no solamente en el marco de la formación que ha seguido o de los empleos que ha ocupado, sino también en el marco de la vida social o personal.

A finales de los años ochenta se implanta en Inglaterra el *National Vocational Qualification* (NVQs) con el fin de intentar aumentar el nivel de cualificación de los trabajadores británicos. Inglaterra ha estado y sigue estando muy interesada en acelerar el reconocimiento y acreditación de este tipo de competencias y la puesta en marcha de dispositivos de reconocimiento con el afán de poder elevar el grado de cualificación de su población trabajadora, considerada por los mismos ingleses como excesivamente baja.

Pero la experiencia, en cuanto a déficit de cualificación de la población trabajadora, no es exclusiva de Inglaterra sino que se extiende al resto de los países europeos en mayores o menores proporciones. La población activa mayor de 24 años que no dispone de cualificación profesional alcanza cifras excesivamente altas en toda Europa. La OCDE declara que en la Europa actual un 60% de los trabajadores no están suficientemente titulados (Werquin, 2007) y, según la Encuesta de Población Activa de la UE 2004 (Eurostat), en Europa hay alrededor de 72 millones de trabajadores poco cualificados, es decir, un tercio de la mano de obra (Comisión Europea, 2006). Por este motivo, productivo fundamentalmente, la Unión Europea, desde la década de los años noventa, está impulsando la creación de nuevos sistemas de reconocimiento y validación de los aprendizajes adquiridos en la experiencia laboral.

En realidad todos los procedimientos de valoración, reconocimiento y acreditación de los aprendizajes no formales o informales nacieron, de una u otra forma, dentro del

ámbito del trabajo, del empleo y de la formación profesional. De ahí que las distintas tradiciones formativas de los distintos países tengan su reflejo ahora en los diferentes sistemas de reconocimiento y acreditación de aprendizajes no formales e informales.

La Unión Europea fomenta el desarrollo de sistemas de reconocimiento de aprendizajes no formales desde mediados de los años noventa. En el año 1996 se publica en el libro blanco sobre «Enseñar y aprender. Hacia la sociedad del conocimiento» y en él se afirma que en la mayor parte de los países europeos el aprendizaje que tiene lugar externamente a la enseñanza formal sigue pasando desapercibido; por ello la Comisión Europea cuestiona sus posibles consecuencias en los siguiente términos: «Podría considerarse que la sociedad “da cerrojazo” a numerosos talentos que con frecuencia resultan no convencionales pero innovadores, y que produce por tanto una élite que realmente no es representativa del potencial disponible de recursos humanos» (Bjørnåvold, 1998, p. 66).

En el citado libro blanco la Comisión Europea (1996, pp. 31-33) distingue tres tipos de aprendizajes a tener en cuenta: los conocimientos básicos (conocimientos adquiridos prioritariamente en la escuela), los técnicos (competencias adquiridas frecuentemente en el puesto de trabajo) y los sociales (competencias interpersonales adquiridas, muchas de ellas, en la vida social).

El mensaje esencial de este libro blanco y de una serie de documentos europeos publicados en los últimos años del siglo XX es que los países y las economías precisan una base más amplia de conocimientos que combine equilibradamente competencias básicas, técnicas y sociales. Así se expone en el capítulo del mismo libro titulado *¿Cuáles son las competencias requeridas?*: «En el mundo de hoy, el conocimiento en sentido amplio puede definirse como un cuerpo adquirido de conocimientos fundamentales o técnicos, unido a competencias sociales. Es el equilibrio de estos conocimientos recibidos a través del sistema educativo formal o a través de la familia, del empleo y de diversas fuentes o redes de información, quien compondrá el cuerpo amplio y transferible de conocimientos más favorable para el empleo» (p. 31).

Cuando en el año 2001, en plena sociedad del conocimiento, Europa se propone (Estrategia de Lisboa 2001) ser la región más competitiva del mundo, se hace ya incomprendible, desde el punto de vista productivo, no reconocer innumerables competencias y conocimientos de los ciudadanos europeos solamente por el hecho de que éstas no hayan sido adquiridas en los sistemas de formación reglada.

Pero en Europa la aplicación y creación de sistemas de reconocimiento y validación ha seguido diferentes caminos y modelos según países o regiones (Comisión Europea, 2001). Esto se debe en gran medida, como ya hemos dicho, a la diferenciación

histórica de sistemas nacionales de formación profesional y a diferentes mentalidades sobre formación profesional. Veamos algunos modelos.

El modelo Francés es uno de los más avanzados y madrugadores en reconocer y validar los aprendizajes adquiridos en la experiencia o en el trabajo, con importantes avances en los métodos y protocolos (Urso, 2005; Zeller, 2006). Durante los últimos veinte años el tema de los aprendizajes no formales ha ocupado un puesto privilegiado en el debate francés sobre enseñanza y formación. Además de los primeros ensayos que comienzan en los años treinta y que se suceden con la legislación de 1985 y de 1992, en el momento actual está en aplicación la denominada Ley de Modernización Social, que reconoce el derecho a ser acreditado (Sanz Fernández, 2006). Conforme a esta ley, cualquier ciudadano puede, a través del sistema de validación, acceder no sólo a la validación parcial de certificados o títulos profesionales sino a la obtención completa del título, incluidos los títulos universitarios, salvo los que están especialmente regulados, como los relacionados con la sanidad o la seguridad (Charraud, 2002).

En Alemania el debate sobre reconocimiento y acreditación de aprendizajes no formales ha comenzado tarde y con poco interés, con una demanda de evaluación de aprendizajes informales escasa (Bjørnåvold, 2002). Alemania, conforme a su modelo de formación profesional, ha considerado siempre que la posibilidad de aprender a partir de la experiencia en el trabajo ya está incluida en su propio sistema de formación dual, que posibilita la participación formativa de las empresas en los mismos lugares de trabajo.

Los países escandinavos tienen experiencias antiguas de validación pero las más importantes datan de los primeros noventa. En Noruega, a partir de 1994 se introducen los aprendizajes adquiridos en la experiencia como parte obligatoria e integrada de todos los cursos. En el mismo año 1994, en Finlandia, una nueva ley sobre Formación Profesional convierte en elemento clave del sistema el que puedan demostrarse las competencias y los conocimientos y reconocerse éstas independientemente de la forma en que se hayan obtenido. En Dinamarca, la Ley de formación para el mercado de trabajo, de 1995, establece con mucha claridad el papel de la formación por experiencia laboral.

El sistema británico de las cualificaciones profesionales (NVQs) ha sido uno de los referentes más claros de formación profesional basada en competencias. Es un sistema que favorece la formación adquirida en la experiencia laboral y considera que lo verdaderamente importante es lo que se ha aprendido y no dónde o cómo se haya aprendido. En opinión de Bjørnåvold (2002), el sistema británico de las NVQs ha

despertado un fuerte interés y se ha convertido, a pesar de las críticas interiores, en un importante punto de referencia a escala internacional.

Los países mediterráneos (Grecia, Italia, España) se han caracterizado por una debilidad de sus sistemas de formación profesional y por la necesidad de promover el aprendizaje de competencias profesionales fuera del sistema reglado. Esto es lo que ha provocado que en estos últimos años, estos tres países hayan acogido con interés todo lo referente a formación profesional, incluyendo el sistema de reconocimiento y validación (Bjørnåvold, 2002).

En el caso concreto de España, con la incorporación a la Universidad de los mayores de 25 años en el año 1970, se inician las primeras pasarelas, que en un sentido amplio podrían asociarse a los sistemas de acreditación. No obstante, dado que hablamos de cuestiones que cuentan con un fuerte ordenamiento normativo, la acreditación de competencias propiamente dicha se inicia a partir del 2002 con la publicación de la Ley Orgánica 5/2002 de las Cualificaciones y de la Formación Profesional, que en su artículo ocho regula todo lo referente al reconocimiento, evaluación y acreditación de competencias profesionales adquiridas a través de la experiencia laboral o mediante vías no formales. A partir de esta primera norma legal se publican otras leyes orgánicas que regulan estas prácticas en algunos sectores educativos: la Ley de Empleo de 2003 (art. 25.2) en su referencia a los programas de formación profesional ocupacional y continua, la Ley de Educación de 2006, que en el art. 66.4 habla de los aprendizajes que obtienen las personas adultas a través de su experiencia y la Ley de Modificación de la LOU de 2007 donde se ponen las primeras bases (art. 36) para validar a efectos académicos la experiencia laboral o profesional en el ámbito universitario.

## **Funciones sociales y educativas de los sistemas de acreditación**

¿Por qué estamos inmersos en este debate? ¿Por qué prácticamente todos los países, y especialmente los de la Unión Europea, están desarrollando estos procesos de reconocimiento y validación de la experiencia? En una primera consulta a la literatura existente, encontramos razones que tienen que ver con los beneficios que obtiene la propia persona que lo solicita, las ventajas que estos procesos tienen para las empresas y los efectos sobre el sistema educativo y la sociedad en general.

## Beneficios para el trabajador

Los que se oponen a los sistemas de acreditación argumentan que el título no añade nada a la competencia y que solo tiene un valor instrumental, no un valor en sí mismo; consideran que lo que hay en el fondo es una mera preocupación credencialista.

Es cierto que algunos autores (Passeron, 1983 y Collins, 1989), desde finales de los setenta, vienen hablando del «credencialismo» como un problema de nuestra sociedad, entendiéndolo por ello la importancia desmesurada que se le da a los títulos académicos y la carrera desenfrenada por adquirirlos; un fenómeno social que ha merecido otras denominaciones como la «fiebre de los títulos», la «inflación de los diplomas», «la sobrecualificación» o la más común de «titulitis» (Fernández Enguita, 1998).

Independientemente del juicio que nos merezca el mero credencialismo, no podemos olvidar la función social que tienen los títulos. Sobre ello, también han escrito no pocos autores (Schultz, 1968; Romero y De Miguel, 1969; Brooking, 1997, entre otros) hablando de las posiciones sociales que se derivan de la titulación y del reconocimiento social que supone poseer un título.

De hecho, reconocen los sociólogos la estrecha relación existente entre educación, distribución de posiciones sociales y acceso al trabajo. No en vano existen evidencias que ponen de manifiesto no sólo que la probabilidad de encontrar empleo es superior entre las personas que poseen un mayor nivel de estudios, sino también la estrecha vinculación entre la formación y el mantenimiento del puesto de trabajo (Gil Villa, 1994). En razón de ello algunos autores (Becker, 1983), posicionándose en el debate formación-empleo, destacan que la mejor inversión que cabe hacer es la que afecta al capital humano. En el fondo, esto nos remite al pacto latente existente entre el Estado y los ciudadanos para la consolidación de los estados de bienestar, pacto en función del cual a cambio de impuestos los estados ofrecen iguales oportunidades de educación y trabajo para todos (Bonal, 1998).

Está claro que hay que poseer las competencias, es evidente, pero no podemos ignorar los perjuicios que se derivan de la ausencia de titulación entre las personas competentes. El problema de los competentes no titulados es sobre todo la ausencia de reconocimiento social, el vacío que le hacen sus colegas y el encontrarse en una especie de limbo profesional de inciertas consecuencias. El tema tiene que ver con el concepto de profesión y sus implicaciones sociales y en la actualidad el título funciona como un medio para evitar ser excluido.

Por tanto, los beneficios de los sistemas de acreditación para la propia persona parecen evidentes. El primero es el que acabamos de comentar: el reconocimiento

social que nace de la titulación y las consecuencias que de ello se derivan en la relación con los propios colegas, la empresa y la sociedad en general.

No menos importante es el beneficio asociado al propio puesto de trabajo, toda vez que, en posesión de un título, el trabajador tiene más probabilidades de mantener su empleo, de promocionar y, sobre todo, de mejorar su salario. Ello es así porque mediante los sistemas de acreditación el trabajador incrementa su propio capital cultural y con ello su valor de cambio en el mercado laboral.

En tercer lugar, la posibilidad de acreditar competencias constituye para el interesado una segunda oportunidad para adquirir una titulación que no consiguió en su día, bien porque no asistió a la escuela, bien porque fracasó en ella, o bien porque no tuvo más remedio que ponerse a trabajar cuanto antes. En este sentido hay que reconocer la función compensatoria que tienen los sistemas de reconocimiento.

Desde el punto de vista personal hay otro beneficio importante: la motivación que crean los sistemas de acreditación entre quienes participan. Muchos trabajadores no entienden (y con razón) que habiendo realizado muchos cursos y actividades formativas a lo largo de su vida, esto no les sirva para la adquisición de un título oficial, que no pudieron conseguir en su día. El resultado es la desmotivación para seguir formándose e integrarse en el aprendizaje a lo largo de la vida.

La experiencia en este sentido nos demuestra dos cosas relacionadas entre sí: una, que si la acción formativa conduce a alguna titulación se incrementa el interés entre los participantes: dos, que si el sistema educativo reconoce lo que se ha adquirido a través de la experiencia se crea entre los candidatos una motivación general para seguir formándose.

## **Beneficios para la empresa**

Los hay que piensan que las empresas no necesitan titulados sino trabajadores competentes, sean o no titulados. Es más, algunas empresas ni siquiera buscan trabajadores que posean unas competencias específicas, porque esperan formar a sus propios trabajadores. Todo parece indicar que, quienes así piensan, en los últimos años se han quedado sin argumentos. De todos es sabido que las empresas, sobre todo, desean contar con trabajadores competentes, no cabe duda, pero también necesitan trabajadores titulados, porque tener trabajadores titulados tiene varias ventajas.

Tal vez la ventaja más reconocida por los diferentes autores es la que aporta evidencias sobre el modo en que el mercado laboral y el crecimiento de la economía se

ajustan a las credenciales educativas de los trabajadores, una cuestión planteada por Schultz (1968) en la teoría del capital humano.

Con todo, independientemente de las críticas que ha recibido Schultz, como la de Thurow, versiones más remozadas de la teoría del capital humano no niegan la relación, al menos indirecta, entre productividad y formación (Gil Villa, 1994). Nos referimos a unas versiones que ponen el énfasis más en cuestiones cualitativas que cuantitativas (la teoría de Schultz subrayó en demasía los elementos cuantitativos). Ello explica, como nos dice Bonal (1998), que la mayoría de las reformas educativas de los países, desarrollados o no, y buena parte de los informes de la OCDE (1991), recuperan en sus textos el valor económico de la educación. También hay que señalar en este sentido las aportaciones de los teóricos de la sociedad de la información (Castells, 1994), que nuevamente ponen de manifiesto la vinculación entre el dominio de las nuevas tecnologías y el desarrollo económico de los pueblos.

La novedad, en todo caso, está en afirmar también la existencia de una cierta relación entre productividad y experiencia, razón por la cual la mayoría de los documentos de las instituciones internacionales (Delors, 1996; Comisión Europea, 2001; OIT, 2004; UGT, 1998, entre otros), insiste en la ventaja que tienen los sistemas de acreditación en lo que se refiere a la transparencia del mercado laboral. Los sistemas de reconocimiento y acreditación, se dice en tales informes, mejoran la transparencia del mercado laboral. Con ello, se quiere decir que un mercado más transparente cuenta con un mayor grado de igualdad de oportunidades en la medida en que se dispone de toda la información necesaria y relevante en los procesos de oferta y demanda de empleo. Dicho de otro modo, un mercado laboral más transparente es un mercado menos vulnerable para los propios trabajadores y menos generador de precariedad y de situaciones de exclusión social. Esto significa, entre otras, dos cosas importantes: por un lado, que las competencias, en la medida en que se reconocen y certifican, se conocen, es decir, se hacen visibles al mercado, a los empleadores y a los propios trabajadores, que ahora saben qué valor tienen en el mercado laboral; y en segundo lugar, que la acreditación facilita la movilidad de los trabajadores y la unidad del mercado laboral, sobre todo en los países de la Unión Europea, que cuentan con titulaciones equivalentes.

Por otra parte, contar con trabajadores titulados, además de competentes, tiene la extraordinaria ventaja de que no sólo permite conocer el capital cultural de la empresa (a través del capital cultural de los trabajadores) sino también incrementarlo, por varias razones: porque el capital cultural es mayor simplemente si «es», es decir, si es visible, reconocido oficialmente; porque podrá ser valorado en los

certificados de calidad que la empresa solicite; y porque va a ser tenido en cuenta en la relación con sus proveedores y clientes. En este caso, los efectos sociales de las titulaciones (de lo que hemos hablado anteriormente) no sólo afectan a los poseedores de los títulos sino a las empresas que cuentan con tales trabajadores titulados. De este modo la empresa podrá mejorar su política empresarial basada en la gestión del capital intelectual de los trabajadores, optimizando así los recursos humanos.

Otra de las ventajas que tienen los sistemas de acreditación para las empresas tiene que ver con la formación de los trabajadores, toda vez que, conocido el capital cultural de los trabajadores y de la empresa, se podrán diseñar y gestionar mejor las acciones formativas que tienen lugar en su seno. No olvidemos que los sistemas de acreditación significan también reconocer la función formativa de la empresa, dado que las competencias adquiridas mediante la formación continua podrán ser acreditadas oficialmente.

Otros documentos tienen más empeño en señalar que la principal ventaja para las empresas de los sistemas de reconocimiento y acreditación es la que se deriva del incremento de la competitividad (Comisión Europea, 2000). Se habla así de las relaciones entre capital humano y competitividad. Es verdad que la competitividad depende tanto del capital físico como del capital humano, pero todo el mundo reconoce que un gran capital físico puede venirse a bajo si no cuenta con el suficiente capital humano (la cultura de los trabajadores, la formación, las competencias, la innovación, el conocimiento, etc.) que lo gestione. Es más, el capital humano puede suplir las debilidades en el capital físico y remover los obstáculos que de ello se deriven, pero no al revés. Ya hemos dicho antes que los sistemas de reconocimiento y acreditación incrementan el capital cultural de los trabajadores y de la empresa, por lo que tales sistemas van a afectar de modo decisivo a la propia competitividad de las organizaciones empresariales.

En relación con los sistemas de acreditación y la competitividad cabe señalar la opinión del Director de Educación de la OCDE en el «Congreso Europeo de Acreditación de Competencias», celebrado recientemente en Oviedo. Decía Werquin (2007) en su conferencia que la crisis económica que sufre Francia desde finales de los noventa se debe en parte también a la incapacidad del mercado laboral francés de dotarse de una mano de obra visible, a partir de finales de los ochenta, una etapa con un importante crecimiento y desarrollo económico. Lo que al mismo tiempo explica el desarrollo que tienen los sistemas de acreditación en Francia a partir del 2003 con la publicación de la Ley de Modernización Social.

## Beneficios para la sociedad en general

Al analizar los beneficios sociales que tienen los sistemas de acreditación tendríamos que remontarnos a sociólogos como Sant-Simon, Comte y sobre todo Durkheim, que aportaron los primeros datos y reflexiones sobre el papel de la educación en la articulación e integración social. Es más, incluso hay aportaciones como la de Dewey, que destacan la estrecha relación entre democracia y educación. Con todo, hay que reconocer que estas primeras posiciones, funcionalistas, contrastan con aproximaciones más críticas, por ejemplo la de Collins (1989) y Bordieu y Passeron (1977), que nos enfrentan al debate sobre el credencialismo y la reproducción.

En cualquier caso, cualquiera que sea la posición de partida (funcionalista o crítica), consideramos que desde el punto de vista social las ventajas de los sistemas de acreditación no son nada desdeñables. En tal sentido, publicaciones más recientes insisten de modo diverso en la misma dirección. Parsons (1990), refiriéndose a la sociedad norteamericana, estudia los procesos selectivos de cara al trabajo que se operan en el mundo de la educación. Planas, por su parte, comenta el valor que tienen los títulos académicos, sobre todo actualmente en la era de la globalización, que operan como filtros en el mercado de trabajo, filtros cada vez más necesarios aunque no siempre suficientes (Planas, 2003). Casal (2003) nos comenta que los procesos de transición de la escuela al trabajo dependen, cada vez más, del valor de cambio de los títulos académicos.

Precisamente por ello, la ventaja que tienen los sistemas de acreditación, que aparece de forma recurrente en los documentos oficiales, es la que habla de tales sistemas como un medio eficaz de promover entre los ciudadanos y trabajadores la cultura del «aprendizaje a lo largo de la vida» (Delors, 1966, p. 160). La razón está en la motivación que crean entre quienes participan y en el interés que despiertan en las empresas (Díaz, 2004).

Asociada a esta ventaja está esta otra, que también encontramos reflejada en documentos internacionales (Delors, 1996, p. 158; Comisión Europea, 1996, p. 39): que por primera vez estamos ante una de las respuestas más efectivas de los sistemas educativos a los problemas sociales, entre otros, la precariedad del mercado laboral, los sectores de población que viven sin trabajo, la población que se encuentra bajo el umbral de la pobreza, etc.

Ello es así porque los sistemas de acreditación no representan otra cosa que la expresión de las políticas de igualdad y de promoción del empleo llevadas a cabo por los gobiernos de los diferentes países, una forma de lucha, en suma, contra las desigualdades y la exclusión social. En este sentido no podemos olvidar las finalidades con las que

nacieron y se consolidaron los sistemas educativos a partir de su generalización en el siglo XIX, entendiendo que la educación es un derecho de todos y que debe convertirse en un sistema para reducir las desigualdades sociales (Comisión Europea, 1996, p. 39). Es más, creemos, como ya sucede en otros países como Francia, que una persona adulta tiene derecho a que se le reconozca lo que ya ha aprendido y que los sistemas educativos tienen el deber de justicia social de atender estas demandas (Medina, 2006, p. 127).

Por otra parte, la apuesta de los gobiernos por los sistemas de acreditación también tiene que ver con lo que ya hemos comentado anteriormente: el hecho de que los sistemas de acreditación de la competencia hacen visible (en tanto que reconocible oficialmente) el capital cultural de los ciudadanos, incidiendo en el incremento general del nivel de estudios (visible, reconocible) de la población.

Finalmente, hemos de comentar otro de los beneficios sociales que tienen los sistemas de acreditación, cual es, dotar de mayor racionalidad a los sistemas educativos. Si una persona tiene las competencias de una determinada profesión, parecería absurdo obligarle a integrarse en el sistema educativo (Lietard, 1992; Aubret, 1991). Hacerlo equivaldría al absurdo de darle más importancia al circuito recorrido (dónde, cuándo y cómo) para adquirir las competencias que a las propias competencias adquiridas, independientemente de cómo, dónde y cuándo se han adquirido. Lo racional en este caso es reconocerle lo que ya sabe, no volverle a enseñar lo que ya ha aprendido con anterioridad. Por ello no son pocos quienes piensan que los sistemas de acreditación aportan más racionalidad y coherencia a los sistemas educativos.

## **Perspectiva teórica de los sistemas de acreditación**

Hasta aquí lo que hemos hecho es presentar algunas bases históricas y socioeducativas sobre los sistemas de acreditación, poniendo de manifiesto su significado y alcance internacional. Pero creemos necesaria una reflexión teórica y pedagógica, sobre todo, para responder a quienes piensan que estamos ante una comprometida aventura pedagógica de inciertos resultados para los sistemas educativos ya consolidados. En nuestro informe de la investigación sobre «Sistemas de Acreditación» abordamos este asunto desde una perspectiva interdisciplinar, tratando de fundamentar los sistemas de acreditación desde la Epistemología, la Historia, la Sociología, la Psicología y la Pedagogía.

No obstante, admitiendo que estamos ante un fenómeno sobre el que cabe una poliédrica lectura teórica y pedagógica, lo que planteamos aquí en este artículo es algo más limitado: por una parte se formulan algunas cuestiones teóricas que nos permiten sin más entrar en el debate (problemas que se plantean, principios que se comparten, perspectiva interdisciplinar); por otra parte, se analizan con algo más de detalle determinadas referencias conceptuales que nos puedan servir para comprender el significado de los sistemas de acreditación en su propio marco teórico.

## Problemas que se plantean

Los problemas e interrogantes que plantean los sistemas de acreditación, que no son pocos ni de menor calado, constituyen la primera cuestión teórica a la que nos enfrentamos (Sanz Fernández, 2006, p. 43-46). Ya hemos comentado que no estamos ante una propuesta rectilínea ni ante una innovación consensuada entre pedagogos y gestores de la educación. El debate y la controversia están servidos y así ha sido casi siempre sobre todo al principio.

La propia Comisión Europea (1996, p. 47) reconoce que estamos ante complejos y diversos debates, por lo que recomienda que intervengan expertos (2000, p. 17), se establezcan redes europeas (1996, p. 61) y se utilicen procedimientos válidos y fiables (1996, p. 35; 2000, p. 18).

Así pues, los problemas son de diverso tipo. Señalamos los que nos parecen más relevantes. Se plantean «problemas pedagógicos», entre otras cosas, porque estamos ante un sistema que se basa en el concepto de competencia, algo sobre lo que no existe suficiente consenso. El problema de la competencia es que se considera un constructo intangible, que no se observa directamente sino a través de conductas; y es precisamente la competencia la que es necesario evidenciar para certificar que una persona es acreedora de una titulación. A partir de aquí se derivan otros problemas: en qué consiste la formación basada en competencia, cómo se elaboran los diseños curriculares a base de competencias, etc.

También tenemos «problemas legales». Cuando hablamos de sistemas de acreditación nos referimos al hecho de reconocer oficialmente competencias que forman parte de una titulación que está regulada por las autoridades competentes; por tanto, son esas mismas autoridades las que deben regular también tales procesos de reconocimiento y acreditación, en la medida en que afectan al modo en que se han regulado las mismas titulaciones. No cabe duda de que estamos ante la necesidad de una

regulación legal de los sistemas de acreditación, pero en este sentido caben varias preguntas: quién tiene la competencia legal para regular los sistemas de acreditación, cómo regular estos sistemas de tal modo que se asegure y no se comprometa el valor social de los títulos, cómo evitar la vía fácil para la titulación, hasta dónde tiene que llegar la normativa y qué funciones asignar a los evaluadores, etc.

No son menores los «problemas sociales» que tienen que ver con la repercusión que tienen estos procedimientos en la población en general y más concretamente entre los educadores profesionales y administradores de la educación. En razón de ello se pueden formular varias preguntas: cómo se justifican y legitiman socialmente los sistemas de acreditación, cómo evitar el rechazo sobre todo en los ambientes académicos, si se trata de una mera opción educativa o estamos ante un nuevo derecho de los ciudadanos, etc.

Con todo, los principales son los «problemas metodológicos». En general los enseñantes estamos acostumbrados a evaluar de lo que previamente hemos enseñado, pero cuando hablamos de sistemas de acreditación se evalúa lo que otros han enseñado o lo que una persona ha podido aprender de forma autónoma o a través de la experiencia. De hecho, quienes se tienen que enfrentar a estos sistemas en la práctica se plantean, entre otras, estas preguntas: cómo se evalúa lo que no se ha enseñado, cómo se evalúan las competencias, qué instrumentos o recursos se usan para la evaluación de la competencia, cuáles son las fases a seguir en la evaluación de la competencia, etc.

Finalmente hay que reconocer que, en el fondo, buena parte de los debates en torno a los sistemas de acreditación nos enfrenta a una serie de «problemas teóricos», de signo filosófico y científico, relacionado con el conocimiento que actualmente tenemos disponible. Los podríamos formular del siguiente modo: ¿se puede aprender sin maestros? ¿La experiencia es fuente de aprendizaje? Si admitimos que la experiencia facilita el aprendizaje de procedimientos, ¿también facilita el aprendizaje de conocimientos (conceptos y teorías) estructurados y jerarquizados?

## **Principios y/o características**

Una segunda consideración teórica es la que responde a la siguiente pregunta: ¿en general qué parece estar cada vez más claro y compartido entre quienes estudian estos temas y entre los profesionales que ya llevan algún tiempo en la práctica de los sistemas de acreditación? Si consultamos las numerosas publicaciones ya existentes sobre sistemas de acreditación (INCUAL, 2003; MECD, 2003; Comisión Europea,

2001; UNESCO, 1985; entre otras), vemos que en general hay acuerdo respecto de una serie de cuestiones que pueden considerarse, o bien principios que rigen esta actividad educativa, o bien características que se dan en la mayoría de los países que lo practican. Algunos de estos principios y/o características los exponemos a continuación:

- Finalidad: el aprendizaje a lo largo de la vida. Tal como se conciben los sistemas de acreditación y tal como se están llevando a cabo en los diferentes países, en el fondo y principalmente, la finalidad no es titular, sino la formación y sobre todo motivar a los trabajadores, promoviendo en ellos una preocupación y una cultura por el aprendizaje a lo largo de la vida.
- Flexibilidad de los sistemas educativos. No parece que pueda hablarse sistemas de acreditación si no existe una cierta flexibilidad en el modo de regular y ordenar los sistemas educativos de los países.
- Regulación legal. En la medida en que los sistemas de acreditación afectan a las titulaciones de las enseñanzas formales, en esa medida deben ser regulados legalmente por las mismas autoridades educativas que tienen las competencias para la ordenación de tales enseñanzas formales.
- Sector de la Educación de Adultos. Hablamos de unas prácticas educativas que tienen sentido en el campo de la Educación de Adultos. No se habla de sistemas de acreditación en las enseñanzas iniciales del régimen general, ni para quienes se encuentran en las edades de escolarización (niños y adolescentes).
- Proceso individualizado. Los sistemas de acreditación no se realizan en grupo. El proceso de acreditación, que fundamentalmente supone la orientación y evaluación de la persona que lo solicita, siempre será un proceso individualizado, mediante la relación directa y personal entre el candidato, el orientador y los evaluadores.
- Funciones diferentes de los orientadores y evaluadores. Esencialmente la acreditación es un proceso de orientación y evaluación para continuar la formación, donde las funciones de los orientadores y de los evaluadores son bien diferentes: el orientador es el que ayuda al candidato a explicitar sus competencias (actúa como la comadrona, o como el abogado defensor del candidato) y el evaluador es el responsable de comprobar si posee o no tales competencias (es el que al final emite un juicio evaluador sobre las competencias del candidato). Más aún, estos dos profesionales de la educación no solo tienen funciones diferentes, sino que a lo largo del proceso de acreditación actúan

con cierta independencia uno del otro, sin perjuicio de que puedan estar coordinados.

- Composición de los evaluadores. Para dar una mayor transparencia al proceso y adecuar la titulación a las exigencias del mercado laboral, en algunos países el equipo de evaluadores está formado por enseñantes de la titulación por la que se solicita ser acreditado, pero también por profesionales de la titulación y por agentes sociales.
- Acreditación parcial y total. Prácticamente en todos los países donde funcionan los sistemas de acreditación se entiende que ésta puede ser parcial o total, según que el proceso termine con la certificación de todas las competencias de una titulación o solo de algunas.
- Solo se valoran las competencias. Un modo de proceder que se está estableciendo entre las instituciones dedicadas a la acreditación es que se distingue claramente entre experiencia y competencias, toda vez que lo que cuenta no son los años de experiencia en sí mismos, ni las horas de formación, ni el itinerario recorrido, sino las competencias (los aprendizajes) que se han podido adquirir durante ese tiempo, que es lo que se evalúa.
- El referente de las competencias. En todo proceso de acreditación siempre existe un referente, que es el documento (currículo, diseño de formación, proyecto educativo) donde consta la relación de competencias, con sus correspondientes realizaciones profesionales, para que los evaluadores comprueben (relacionen) si el candidato las posee.
- Pruebas válidas y fiables. Las normas legales que regulan los sistemas de acreditación suelen insistir en la necesidad de que las pruebas, instrumentos y protocolos que evidencian las competencias tienen que estar diseñadas con el rigor necesario que asegure la fiabilidad (grado de constancia en los resultados) y la validez (que realmente evidencien lo que se pretende evidenciar).
- Justicia y viabilidad. Los sistemas de acreditación tienen que basarse en los principios de justicia social (equidad, transparencia, confianza, credibilidad, legitimidad, calidad, etc.) y viabilidad (laboral, económica, funcional, organizativa, etc.).
- Profesiones no reguladas. Los sistemas de acreditación que conducen a una titulación oficial sólo se aplican a aquellas profesiones que no están altamente reguladas. No se aplican, por ejemplo, para la obtención del título de Maestro o Médico, porque legalmente no se puede trabajar en estos contextos laborales sin la titulación correspondiente.

## Perspectiva interdisciplinar

También podemos aproximarnos a este fenómeno con una reflexión sobre las bases teóricas que podrían justificar los sistemas de acreditación (Sanz Fernández, 2007). En tal sentido, nos parece que lo primero que hay que destacar es que, a partir de la bibliografía que hemos consultado, todo parece indicar que los sistemas de acreditación nacen y se desarrollan en la sociedad de la información por dos razones principales, que constituyen al mismo tiempo las dos grandes finalidades de los sistemas de acreditación de la competencia: en primer lugar para promover y motivar a los trabajadores en todo lo que significa el aprendizaje a lo largo de la vida; y en segundo para luchar de forma más eficaz contra las desigualdades y la exclusión social.

No obstante, independientemente de las razones y finalidades de los sistemas de acreditación, otra cuestión es la justificación y legitimidad que nos merecen estas nuevas formas de titulación, como algunos prefieren denominar. Desde nuestro punto de vista, consideramos que los sistemas de acreditación deben fundamentarse y legitimarse de forma sistemática e interdisciplinar en los siguientes saberes que le van a aportar el basamento teórico de lo que representa en la práctica. Estos saberes son la Epistemología, la Historia, la Psicología, la Sociología y también la Pedagogía.

En primer lugar, sería necesario indagar en la Epistemología sobre el papel de la experiencia y de la práctica como fuentes de conocimiento (conceptos, procedimientos y valores). En tal sentido la tradición filosófica del Empirismo y las aportaciones de la denominada Nueva Epistemología podrían ofrecernos suficientes elementos de reflexión al respecto.

La Historia es otra de las disciplinas que nos permite aproximarnos a este fenómeno a la búsqueda de las huellas y experiencias educativas del pasado para conocer los antecedentes de los sistemas de acreditación.

También es necesario acudir a la Psicología para analizar los conocimientos y teorías disponibles en relación con la relevancia del aprendizaje en la edad adulta y sobre el papel de la experiencia en los procesos de aprendizaje. Hablaríamos en este caso de una reflexión y análisis sobre el aprendizaje experiencial.

No menos importante es la lectura de la Sociología y la literatura social en general a la búsqueda de la relación entre la sociedad de la información y las situaciones de exclusión social. En este marco de análisis pensamos que los sistemas de acreditación juegan un papel clave como medios para reducir las desigualdades en los sistemas educativos.

Finalmente cabe también la reflexión estrictamente Pedagógica como el marco más amplio e integrador de los otros saberes en el que cada vez más tienen lugar los

conocimientos y las teorías que se están desarrollando sobre los sistemas de acreditación de la competencia. Más concretamente pensamos que ese marco más amplio es la Pedagogía Social, toda vez que aquí nos encontramos con ideas fuerza y campos de intervención que están resultando decisivos en las prácticas de acreditación de la competencia: entre otros, nos referimos al concepto de educación permanente o, como ahora se prefiere decir, aprendizaje a lo largo de la vida, y a la Educación de Personas Adultas como un ámbito específico de aplicación.

### **Tipos de acreditación y la inevitable distinción entre educación formal, no formal e informal**

Situados en la reflexión pedagógica, otra línea de aproximación teórica es la que pretende clasificar los sistemas de acreditación y definir el marco conceptual previo en el que se identifican con propiedad. Desde nuestro punto de vista, creemos que cabe alguna clasificación en base a la distinción entre los conceptos de educación formal, no formal e informal. Independientemente del valor que les atribuyamos a estos conceptos y al margen de lo afortunado o desafortunado de tales denominaciones (algunos autores opinan sobre ello: Ortega Esteban, 1998, p. 165; García Carrasco, 1988, p. 152; Melich, 1994, p. 69), pensamos que se trata de una distinción inevitable, por la naturaleza de los elementos que intervienen en los procesos de acreditación.

Si acudimos a la legislación sobre sistemas de acreditación de cualquier país, lo primero que constatamos es que en buena parte de las normas legales subyace un concepto amplio de acreditación. Así vemos que no sólo se admiten pasarelas (que dispensan de los requisitos de algún título), sino también pruebas únicas (un examen, sin requisitos académicos previos, que si se aprueba conduce a la titulación) y lo que consideramos que son sistemas de acreditación propiamente dichos (que, mediante la evaluación de las competencias adquiridas en la experiencia, se puede obtener algún título oficial o certificado). Además se trata de una legislación referida al reconocimiento de lo adquirido que tiene lugar, en la mayoría de los casos, en el ámbito de las enseñanzas formales reguladas por las autoridades educativas, pero también, en algunos casos, en el ámbito de las enseñanzas no formales, reguladas por las autoridades laborales.

No obstante, conviene aclarar que hay otros casos en los que la acreditación no dispensa de nada, ni conduce a ningún título o certificado de profesionalidad. Se da este caso, por ejemplo, cuando una persona solicita ser acreditada, pero el tribunal

considera que no tiene todas las competencias adquiridas que se precisan para un título oficial o para un certificado de profesionalidad. Lo que hace el tribunal es certificar solo determinadas competencias, pero no todas. En estos casos también la acreditación tiene sentido, toda vez que los candidatos han tenido la ocasión de reflexionar y explicitar las competencias que poseen y contar con un dossier donde se detalla todo su recorrido formativo y laboral y donde se describen las competencias que han podido adquirir a lo largo de ese proceso.

Hay otras experiencias de acreditación que no están legalmente reguladas ni tienen por qué estarlo, porque no tienen que ver con las enseñanzas formales o no formales reguladas por las autoridades educativas o laborales. Son experiencias de acreditación que no se asocian a título o certificado alguno, cuyas normas las establecen las propias instituciones en cuyo seno tienen lugar para sus trabajadores y/o usuarios. Un claro ejemplo de ello es el programa CERTICAP de *Universia*: se trata de un sistema universitario de acreditación de competencias transversales en idiomas, ofimática y otras competencias (personales, emprendedoras, de participación, de gestión, de potencial de aprendizaje, etc.), cuya certificación la realiza la propia Universidad que está integrada en el programa (Vázquez Vega, 2007).

En consecuencia, a partir de la legislación de los diferentes países, y considerando otros casos y experiencias existentes, podemos hablar de varios tipos de sistemas de acreditación, necesariamente agrupados en dos categorías según estemos hablando de enseñanzas formales o no formales:

- *Acreditación en el ámbito de las enseñanzas formales.* Se trata de sistemas de acreditación regulados por las autoridades educativas. La acreditación en este ámbito puede ser de tres tipos:
  - Pasarelas que acreditan el acceso. Se realizan con la finalidad de dispensar del requisito de un título o de estudios previos para seguir (acceder a) determinadas enseñanzas. Por ejemplo, el acceso a la Universidad para mayores de 25 años, o la acreditación de las enseñanzas parciales de Formación Profesional.
  - Pruebas únicas que acreditan para un título oficial. Se realizan con la finalidad de obtener un título mediante una única prueba o examen; por ejemplo, las pruebas únicas que se establecen en la LOGSE, la LOCE y la LOE para los títulos de Graduado en Educación Secundaria, Bachillerato y Formación Profesional.
  - Sistemas de acreditación propiamente dichos que acreditan para un título oficial. Se realizan con la finalidad de obtener un título, o parte del mismo,

por medio de un sistema específico de evaluación de competencias, las cuales se han podido adquirir mediante la experiencia (educación informal) o mediante enseñanzas no formales. Ejemplos: la acreditación de la formación básica de adultos de Canarias, o la acreditación de competencias para los títulos de Técnico y Técnico Superior de Formación Profesional.

■ *Acreditación en el ámbito de las enseñanzas no formales.* Otros sistemas de acreditación tienen lugar en el ámbito de las enseñanzas no formales, por lo que no están regulados por las autoridades educativas. Distinguimos dos casos:

- La acreditación de la formación ocupacional y continua. En su regulación intervienen las autoridades laborales: el Ministerio de Trabajo y/o las Consejerías de Empleo de las Comunidades Autónomas. En estos casos se acredita para un certificado de profesionalidad, o parte del mismo, del sistema de formación ocupacional y/o continua y también se hace mediante un sistema específico de evaluación de la competencia; por ejemplo, las acreditaciones que están realizando actualmente el propio Ministerio de Trabajo y los Institutos Regionales de las Cualificaciones Profesionales.
- La acreditación de competencias de carácter general. No tienen relación con ninguna titulación o certificado y se regula por las propias instituciones que lo desarrollan. Lo que se busca es acreditar una competencia de carácter general que el interesado puede hacer valer en su currículum; por ejemplo, la acreditación de CERTICAP de Universia, comentado anteriormente.

Como vemos por lo dicho hasta ahora, al hablar de reconocimiento de lo adquirido y acreditación es inevitable usar los términos clásicos de *formal*, *no formal* e *informal* referidos a la educación. Ello es así por dos razones: porque, en la mayor parte de los casos, la acreditación consiste en el valor educativo, en términos de enseñanza formal, que damos a las competencias que se han podido adquirir mediante sistemas no formales o informales, y en segundo lugar, porque, como acabamos de ver, los sistemas de acreditación tienen desigual valor y se realizan de modo diferente según tengan lugar en el ámbito de las enseñanzas formales o en el ámbito de las enseñanzas no formales.

Lo cual quiere decir que, para comprender mejor en qué consisten y cómo funcionan los sistemas de acreditación es imprescindible distinguir claramente entre estos conceptos tradicionales de educación formal, no formal e informal.

No vamos a entrar aquí en profundidad en esta distinción, ya que esto lo han hecho convenientemente autores conocidos por todos, entre otros, Coombs (1971), Touriñan (1983), García Carrasco (1988) y Trilla (1993). No obstante, de acuerdo con los autores citados y de cara a nuestro trabajo, sí debemos precisar algunos extremos. En primer lugar, debemos distinguir entre dos grandes categorías: una, la educación formal y no formal, y otra, la educación informal. En segundo lugar, aclarar que tanto la educación formal como la no formal, se refieren a actividades intencionalmente educativas que además se realizan de forma sistemática; por el contrario, hablamos de educación informal cuando se producen cambios educativos per accidens, es decir, a través de actividades que se han podido realizar sin una intencionalidad educativa expresa y de forma no sistemática. Tercero, la diferencia entre educación formal y no formal, depende de la intervención de las autoridades educativas, pues se considera enseñanza formal cuando tales autoridades deciden el currículo obligatorio para un nivel educativo que, además, se refrenda con un título oficial.

Establecidas estas precisiones, conviene aclarar que, cuando hablamos de acreditación formal en sentido estricto, estamos hablando de una acción educativa consistente en: *un juicio evaluador, emitido por profesionales de la educación formal, consistente en identificar, reconocer y certificar determinados aprendizajes, adquiridos mediante sistemas no formales e informales, que equivalen a las competencias de una titulación oficial.*

## Referencias bibliográficas

- AUBRET, J. (1991). Rédiger un portefeuille de compétences: se reconnaître pour se faire reconnaître. *L'orientation scolaire et professionnelle*, vol. 20, 1. En LIETARD, *op. cit.*
- AUBRET, J. ET AL. (1989). Reconnaissance des acquis et outils d'évaluation. *Pratiques de Formation* (Collection Thématiques) Mars 1989. Université de Paris VIII.
- (1992). Reconnaissance et validation des acquis: le portefeuille de compétences. *Pratiques de Formation* (Collection Thématiques) Septiembre 1992. Université de Paris VIII.
- BEAUDET, A. (1984). La reconnaissance des apprentissages expérientiels: une approche selon les besoins de groupes particuliers de population. *Prospectives*, 20, 3., 157-158.
- BECKER, G. S. (1983). *El capital humano*. Madrid: Alianza.
- BJØRNÅVOLD, J. (1998). La evaluación del aprendizaje no formal: calidad y limitaciones de las metodologías. *Formación Profesional*, CEDEFOP, 12, 58-75.

- BONAL, X. (1998). *Sociología de la educación. Una aproximación crítica a las corrientes contemporáneas*. Barcelona: Paidós.
- BORDIEU, P Y PASSERON, J. C. (1977). *La reproducción. Elementos para una teoría del sistema de enseñanza*. Barcelona: Laia.
- BROOKING, A. (1997). *El capital intelectual*. Barcelona: Paidós.
- CASAL, J. (2003). La transición de la escuela al trabajo. En F. FERNÁNDEZ PALOMARES (Coord.), *Sociología de la educación* (179-201). Madrid: Pearson Educación.
- CASTELLS, M. (1994). Flujos, redes e identidades: una teoría crítica de la sociedad informacional. En M. CASTELLS, R. FLECHA, P. FREIRE, y otros, *Nuevas perspectivas críticas en educación* (15-53). Barcelona: Paidós.
- COLLINS, R. (1989). *La sociedad credencialista*. Madrid: Akal. [1979].
- COMISIÓN EUROPEA (1996). *Enseñar y aprender. Hacia la sociedad del conocimiento. Libro Blanco*. Luxemburgo: Oficina de Publicaciones Oficiales de las Comunidades Europeas.
- (2000). *Memorandum sobre el aprendizaje permanente*. Bruselas: Comisión de las Comunidades Europeas.
- (2001). *Iniciativas nacionales para promover el aprendizaje a lo largo de la vida*. CEDEFOP Y EURIDICE. Madrid: Subdirección General de Información y Publicaciones del MEC.
- (2006). *Aprendizaje de adultos: Nunca es demasiado tarde para aprender*. Bruselas: Comunicación de la Comisión Europea.
- COOMBS, F. (1971). *La crisis mundial de la educación*. Barcelona: Península.
- DELORS, J. ET AL (1996). *La educación encierra un tesoro*. Madrid: Santillana-UNESCO.
- DÍAZ, T. (2004). Certificación de competencias en el marco de la formación continua: contexto y metodología. *Diálogos. Educación y Formación de Personas Adultas*, 39-40, 7-12.
- FERNÁNDEZ ENGUITA, M. (1998). Credencialismo. En S. GINER Y E. LAMO DE ESPINOSA (Dir.). *Diccionario de Sociología* (p. 158). Madrid: Alianza.
- GARCÍA CARRASCO, J. (1988). *Agentes de la educación formal, no formal e informal*. En Symposium Internacional de Filosofía de la Educación (127-155). Barcelona, Bellaterra.
- GIL VILLA, F. (1994). *Teoría sociológica de la educación*. Salamanca: Amaru.
- INCUAL (2003). *Sistemas Nacionales de Cualificaciones y Formación Profesional*. Madrid: Instituto Nacional de Empleo.
- JEAN, M. (1982). *Apprendre: une action volontaire et responsable: énoncé d'une politique globale de l'éducation des adultes dans une perspective d'éducation permanente*. Québec: Ministère des Communications.

- LIÉTARD, B. (1992). Institutionnalisation d'une pratique: les bilans. *Education Permanente*, 108, 137-146.
- MECD (2003). *Proyecto experimental para la evaluación, reconocimiento y acreditación de las competencias profesionales* (Proyecto ERA). Madrid: Ministerio de Educación Cultura y DeportesAutor.
- MEDINA, O. (2006). Los sistemas de acreditación. Aproximación conceptual y teórica. *Educar*, 38, 105-131.
- MÉLICH, J. C. (1994). *Del extraño al cómplice. La educación en la vida cotidiana*. Barcelona:Anthropos.
- MEYER, N. *et al.* (1987). Expériences nord-américaines en reconnaissance des acquis (Québec-Etats Unis). *Pratiques de Formation* (Numéro Spécial) Avril 1987. Université de Paris VIII.
- OTI (2004). Resolución 195, de junio de 2004 sobre la formación y el aprendizaje permanente. Ginebra: Organización Internacional del Trabajo.
- ORTEGA ESTEBAN, J. (1998). La educación a lo largo de la vida o el espacio de la educación social. En PANTOJA, L. (Ed.), *Nuevos espacios de la educación social* (161-182). Bilbao: Deusto.
- PARSONS, T. (1990). El aula como microsistema social. Sus funciones en la sociedad americana. *Educación y Sociedad*, 6, 173-195.
- PASSERON, J. C. (1983). La inflación de títulos escolares en el mercado de trabajo y el mercado de los bienes simbólicos. *Educación y Sociedad*, núm. 1.
- PLANAS, J. (2003). Educación y mercado de trabajo en la globalización. En F. FERNÁNDEZ PALOMARES (Coord.). *Sociología de la educación* (165-178). Madrid: Pearson Educación.
- ROMERO, J. L. Y DE MIGUEL, A. (1969). *El capital humano*. Madrid: Confederación de Cajas de Ahorros.
- SANZ FERNÁNDEZ, F. (2006). Reconocimiento y validación de los aprendizajes adquiridos en la experiencia: la ley francesa de modernización social. En LARA GUIJARO Y QUINTANAL DÍAZ (Coords.). *El practicum en las titulaciones de educación: reflexiones y experiencias* (pp. 31-47). Madrid: Dykinson.
- (2007). Pensar la relación entre experiencia y aprendizaje. *Miscelánea Comillas*, 65, 126, 237-265.
- SCHULTZ, T. W. (1968). *Valor económico de la educación*. México: Uteha.
- TOURIÑÁN, J. (1983). Análisis teórico del carácter formal, no formal e informal de la educación. *Papers d'Educació*, 1, 105-127.
- TRILLA, J. (1993). *La educación fuera de la escuela*. Barcelona: Ariel.

- Ugt (1998). *Informe comparativo de modelos europeos de acreditación de competencia*. Proyecto Valprof.
- U+UNESCO (1985). *Cuarta Conferencia Internacional de Educación de Adultos. Informe final*. París: UNESCO.
- URSO, J. P. (2005). *Comment valider les acquis de son expérience*. Paris: Chiron.
- VÁZQUEZ VEGA, A. (2007). *Certicap: Presentación de una Metodología de Certificación de Conocimientos, para Universitarios. Actas del Congreso Europeo de Acreditación de Competencias*. Oviedo: Cámara de Comercio.
- WERQUIN, P. (2007). *Recognition of Non Formal and Informal Learning in the OECD: Emerging Issues*. Actas del Congreso Europeo de Acreditación de Competencias. Oviedo: Cámara de Comercio.
- ZELLER, O. (2006). *Valider ses acquis*. Paris: Groupe Express.

## Fuentes electrónicas

- BJØRNÁVOLD, J. (2002). *Identificación, evaluación y reconocimiento de los aprendizajes no formales: tendencias europeas*. En Agora V. Detección, evaluación y reconocimiento de competencias no formalizadas (9-36). CEDEFOP. Panorama series 46. Luxemburgo: Oficina de Publicaciones Oficiales de las Comunidades Europeas. Consultado el 15 de mayo de 2007, de:  
[http://www2.trainingvillage.gr/etv/publication/download/panorama/5132\\_ES.pdf](http://www2.trainingvillage.gr/etv/publication/download/panorama/5132_ES.pdf)
- CHARRAUD, A. M. (2002). *El reconocimiento y la validación de competencias informales en Francia*. En Agora V. Detección, evaluación y reconocimiento de competencias no formalizadas (115-123). Cedefop. Panorama series 46. Luxemburgo: Oficina de Publicaciones Oficiales de las Comunidades Europeas. Consultado el 15 de mayo de 2007, de:  
[http://www2.trainingvillage.gr/etv/publication/download/panorama/5132\\_ES.pdf](http://www2.trainingvillage.gr/etv/publication/download/panorama/5132_ES.pdf)

**Dirección de contacto:** Óscar Medina Fernández. Universidad de las Palmas de Gran Canaria. Facultad de Formación del Profesorado. c/ Juana de Arco, 1. 35003 Las Palmas de Gran Canaria, España. E.mail: omedina@dedu.ulpgc.es



# La nueva reforma educativa argentina según sus bases legales

## The Argentinean new educational reform analyzed from its legal framework

Guillermo Ruiz

*Universidad de Buenos Aires, Argentina. Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)*

### Resumen

En este artículo se analiza la legislación educativa que se sancionó en la Argentina a partir del mes de agosto de 2005, durante el gobierno del Presidente Kirchner. Para ello se considerarán como unidades de análisis las leyes y proyectos de leyes presentados por el gobierno nacional. Se realiza un encuadre histórico comparativo con la legislación sancionada en los años noventa, la cual dio lugar a una reforma estructural del sistema educativo nacional. Se describen los cambios entonces efectuados en la estructura académica y la organización institucional del sistema.

La actual reforma está constituida, en principio, por dos leyes sancionadas durante el año 2005 que regulan aspectos parciales de la educación, y una nueva ley orgánica para la educación argentina sancionada a fines del año 2006. Por esta última se derogó la legislación de los años noventa y se estableció una modificación en la estructura académica que se tornará problemática de aplicar porque resulta también problemático el punto de partida en este aspecto debido a la desigualdad educativa que prevalece en el sistema educativo nacional.

Resulta necesario aprender de las lecciones que dejó la aplicación de la Ley Federal de Educación de 1993. Si no se consideran las condiciones en las que se desarrolla la enseñanza en la Argentina como consecuencia de aquella ley, es poco probable que se alcancen algunos de los objetivos fijados en este proyecto de ley. Pero nuevamente aparece la falta de previsión de las consecuencias que tendría una nueva modificación de la estructura académica, en el

marco de niveles de gobierno con competencias concurrentes para aplicarlas de forma diferenciada. Ello constituye un problema muy serio cuyas consecuencias parecen no ser tenidas adecuadamente en cuenta.

*Palabras claves:* reforma educativa, sistema educacional, estructura académica, política educativa, leyes educativas.

### **Abstract**

This article focuses on the Argentinean new educational reform that was approved in Argentina from August 2005, during the government of President Kirchner. The units of analysis to be used are the educational laws and bills designed by the current national government. A comparative and historical analysis is made of the approved '90s legislation, which led to a structural reform of the educational system. A description is given of the changes then applied to both its academic structure and the institutional organization of the system.

In principle, the current reform consists of two national laws, authorised in 2005, which regulate partially educational issues and a new national law for education, which was introduced and passed by the Parliament at the end of 2006. The legislation from the '90s is countermanded by this new 2006 education law, which establishes a new academic structure. The implementation of such a structure could be challenging given that, in that respect, the starting point is also difficult due to the inequality and discrimination in the national education system.

It is necessary to learn from the 1993 Federal Law on Education. If we do not take notice of the conditions in which Argentinean education occurs, as a consequence of the implementation of the afore-mentioned law, the achievement of the objectives set by this bill is highly unlikely. We can see the emergence once again of the absolute lack of foresight of the potential consequences of further alterations in the academic educational structure, in the scope of government authorities holding concurrent competences whilst trying to implement the new structure differently. It constitutes a serious problem, the consequences of which seem not to be sufficiently considered.

*Key Words:* educational reform, educational system, academic structure, educational policy, educational laws

## Introducción

En este artículo se describen las acciones legislativas que lleva adelante el gobierno del Presidente Kirchner (2003-07), las que terminarán por modificar la organización institucional y la estructura académica del sistema educativo argentino. Para ello se toma como unidad de análisis a las leyes educativas sancionadas por el Congreso Nacional durante este gobierno ya que se considera, siguiendo a Paviglianiti, que la legislación puede ser entendida como la expresión normativa del diseño y ejecución de las políticas educativas por parte del estado.<sup>1</sup> De todos modos, para enriquecer esta conceptualización aquí se concibe a la legislación como parte de las *políticas públicas* ya que esto incluye a todas las variantes empíricas de programas de acción (leyes, resoluciones) establecidos por las burocracias públicas. Ello comprende distintos niveles y grados de articulación de objetivos y especificaciones para instrumentar regulaciones normativas. En la concreción de toda política pública intervienen diversos actores, cuyos conflictos de intereses, en cuanto a los contenidos normativos de una política, afectan el proceso entero de concreción de ésta. Así, la relación Estado-Sociedad se estructura en torno a las diversas *tomas de posición o políticas* de los diferentes actores sociales involucrados y afectados. El aparato estatal interioriza nuevas tensiones al privilegiar o censurar determinados intereses sectoriales en detrimento de otros.<sup>2</sup>

En consecuencia, se estudia la política educativa del gobierno de Kirchner a través de las normas destinadas a reformar el sistema educativo. Para ello se analiza la lógica de propósitos enunciados en relación con los contenidos de las normas educativas así como la coherencia interna de cada instancia de estas normas entre sí. En el período comprendido entre agosto de 2005 y diciembre de 2006 el gobierno argentino diseñó una serie de proyectos que se transformaron en leyes que dieron lugar a una nueva reforma educativa.

## La legislación educativa sancionada en los años noventa

La Argentina llevó adelante una reforma estructural de su sistema educativo durante los años noventa. Ésta debe ser entendida como parte, no sólo del proceso de reforma

<sup>(1)</sup> PAVIGLIANITI, N. (1988). *Diagnóstico de la administración central*. Buenos Aires: ME y J.

<sup>(2)</sup> OSZLAK, O. (1992). *Diagnóstico estratégico para la reforma del Estado en América Latina: orientaciones, metodologías e instrumentos*. Buenos Aires: Mimeo.

del Estado, sino también como un *eslabón* en la cadena de continuidades evidenciadas desde mediados del siglo XX, por las cuales el Estado Nacional ha adoptado un papel subsidiario en la prestación del servicio educativo, vía la transferencia de establecimientos educativos nacionales de nivel primario a los gobiernos jurisdiccionales y la desregulación de la educación privada. Resulta necesario destacar lo dispuesto en el art. 14 de la Constitución Nacional, donde se garantiza el *derecho a enseñar y a aprender* a todos los habitantes de la Argentina. En virtud de ello se puede establecer el papel que asume el Estado para garantizar este derecho si se analizan las políticas educativas instrumentadas en cada contexto histórico. Así, cuando se privilegia el *derecho a aprender* el Estado tendría un *papel principal en materia educativa*; este derecho constituye uno de los derechos fundamentales del ser humano, supone el acceso a los niveles más elevados del conocimiento. Desde esta concepción el Estado debe proveer recursos (salariales, edilicios, didácticos, bibliográficos) a la población para garantizar el ejercicio del *derecho a la educación*, tal como comenzó a denominarse desde mediados del siglo XX a este derecho gracias a la conceptualización de los *derechos sociales*. Al contrario, cuando desde las políticas educativas prevalecen las posturas que enfatizan el *derecho a enseñar* (sobre el de aprender) el Estado adoptaría un *papel subsidiario respecto de otros agentes de la sociedad* ya que promovería en primer lugar el derecho original a educar que le corresponde a los padres, a las iglesias y los particulares. Según estas posturas, el Estado debería intervenir sólo cuando la actividad educativa privada no resulte suficiente para cubrir las necesidades de la población.

La ley 24.049 constituyó un punto de llegada para un proceso histórico en las políticas educativas argentinas, que se inició erráticamente en los años sesenta y que concluyó con la rápida sanción de esta norma a fines de 1991<sup>3</sup>. Dicha celeridad quedó evidenciada desde el inicio de esta política al sancionarse esta ley el seis de diciembre de 1991 (Boletín Oficial 07/01/92). El mecanismo para efectuar la transferencia fue el establecimiento de convenios específicos entre cada jurisdicción provincial y el Poder Ejecutivo Nacional (PEN). Es una norma escueta, posee veintisiete artículos, distribuidos en siete capítulos. No formaba parte ni era la especificación de una ley marco para el conjunto del sistema educativo. Al contrario, la ley pareció responder a las necesidades de recorte presupuestario de los gastos del gobierno

<sup>3</sup> La Argentina es un estado federal, constituido por 23 estados federales, denominados *provincias* las cuales son autónomas a las que se suma el Gobierno de Ciudad Autónoma de Buenos Aires (GCBA) que es la capital del país. En este artículo se hará referencia a las provincias y al GCBA indistintamente como «jurisdicciones provinciales», «estados federados» o «estados provinciales».

nacional al transferir a las jurisdicciones provinciales los servicios educativos sin pautas pedagógicas específicas.

Por medio de una vaga terminología la norma propugnó una cobertura educativa basada en los conceptos de *calidad pedagógica* y *equidad educativa* entre los distintos estados federados pero sin especificar los mecanismos que permitirían el cumplimiento de dichos objetivos. Sin embargo, la transferencia de los servicios fue muy criticada por los gremios docentes y por la comunidad educativa a través de diversas manifestaciones de protesta. Incluso los propios gobiernos provinciales condicionaron la firma de los convenios con el Estado nacional a la existencia de partidas presupuestarias que les permitieran hacer frente a los costos que esta política conllevaba para sus administraciones jurisdiccionales. La actitud del gobierno nacional fue intransigente y al cabo de dos años se completó la transferencia, sin estudios diagnósticos que estimaran sus consecuencias económicas, políticas y pedagógicas.

La ley dejó vacíos normativos. No avanzó en materia relativa al régimen laboral que regularía a los docentes después de la transferencia, ni tampoco dio pautas para la solución de las incongruencias entre las normas nacionales y jurisdiccionales. Solamente en el art. 13º, referido a los concursos pendientes y a los procedimientos vinculados con las titularizaciones, traslados y jubilaciones, estableció que éstos debían «ser resueltos según la normativa de origen con el alcance y en tiempos y modos que establezcan los respectivos convenios bilaterales».

Un artículo de esta ley 24.049 que merece destacarse es el 23º:

«Los servicios educativos de gestión privada que se transfieren quedan garantizados para que se sigan prestando con *respeto de los principios* de la libertad de enseñanza y *los derechos emergentes de la normativa nacional* sobre la materia. Consecuentemente, *podrán mantener sus características* doctrinarias, *modalidades curriculares y pedagógicas* y el estilo ético formativo propios, en el ámbito de las jurisdicciones receptoras. Estas *mantendrán el régimen de aportes* a la enseñanza privada en concordancia con el  *fijado en el orden nacional*, en lo que hace a montos, proporciones y destino a los servicios transferidos». (Lo destacado es propio)

Como puede observarse, si bien la ley resultaba imprecisa en general es muy clara sobre los derechos que se le garantizaban a la educación privada transferida. Habría que considerar además el contexto económico de la transferencia educativa, signado por los recortes presupuestarios y el retroceso del Estado en el sostenimiento de los servicios sociales. Así, la política de *descentralización educativa* instrumentada por el gobierno del Presidente Menem respondió al predominio ideológico de concepciones

propias de la nueva derecha<sup>4</sup> que al igual que en otros países promovieron nuevas formas organizativas para la educación pública.<sup>5</sup> En estas políticas resultó predominante el papel de organismos internacionales como el Banco Mundial y el Banco Interamericano de Desarrollo<sup>6</sup>.

A partir de la sanción de esta ley 24.049, y más aún de la Ley Federal de Educación (sancionada en abril de 1993 como ley 24.195), se estableció así un esquema de gobierno y administración del sistema educativo nacional basado sobre la acción directa de las jurisdicciones provinciales y la acción indirecta (y focalizada, esto es, a instituciones o programas en particular y no a las jurisdicciones en su conjunto) del Estado Nacional. Esta acción concurrente se evidencia en el hecho de que el Estado Nacional conserva la facultad de otorgar la validez nacional de los títulos y carreras, lo cual quedó sujeto a la aplicación de la estructura académica diseñada en la Ley Federal de Educación y a los acuerdos alcanzados en el seno del Consejo Federal de Cultura y Educación (CFCyE)<sup>7</sup> a partir de las pautas y criterios generales fijados por la propia ley 24.195.

<sup>4</sup> La nueva derecha engloba concepciones derivadas de una corriente económica denominada *neoliberalismo* por la que se propone una nueva forma de estado que conjuga experiencias de gobiernos *neoconservadores*, tales como los de Margaret Thatcher y John Major en Gran Bretaña, Ronald Reagan en los Estados Unidos y Brian Mulroney en Canadá, con otras propias del neoliberalismo implementado en América Latina como el gobierno inconstitucional del Gral. Augusto Pinochet de Chile. Durante la década de 1990 se identifican los gobiernos de Carlos Saúl Menem en la Argentina y de Carlos Salinas en México como casos de modelos neoliberales. Véase al respecto TORRES & MORROW. El Estado, los movimientos sociales y la reforma educacional. En Mora-Ninici & Ruiz (eds.), *Sociología política de la educación en perspectiva internacional y comparada: las contribuciones de Carlos Alberto Torres*. Buenos Aires: Miño y Dávila, en prensa.

<sup>5</sup> Estas reformas educativas que aparecieron con mucha similitud en los países latinoamericanos durante los años noventa se encuadraron a su vez en los programas de estabilización y ajuste de la economía nacional propugnados por el Banco Mundial y Fondo Monetario Internacional. Dichos programas incluyeron la reducción del gasto del estado, la devaluación de la moneda para promover las exportaciones, la reducción de los aranceles a las importaciones y el incremento del ahorro público y privado. Un aspecto central del modelo es la drástica reducción del sector estatal, especialmente a través de la privatización de las empresas públicas, la liberación de las políticas salariales y de los precios y la reorientación de la producción industrial y agrícola a través de las exportaciones. En el corto plazo, el objetivo de este paquete de políticas es reducir el tamaño del déficit fiscal, de la deuda pública, de la inflación, de las tasas de cambio y de las tarifas. En el largo plazo, el ajuste estructural está basado sobre la premisa de que la exportación constituye la bisagra para el desarrollo.

<sup>6</sup> Estos organismos comenzaron a producir documentos en los cuales se apoyaba la descentralización completa de los sistemas de educación pública y la mayor participación del sector privado en la prestación de los servicios, el otorgamiento de presupuesto asociado al rendimiento y a la *calidad* de la educación así como el subsidio con fondos públicos al sector privado.

Respecto de la influencia que los organismos internacionales han comenzado a tener en las políticas educativas de los países latinoamericanos, véase entre otros: BANCO MUNDIAL (1987). *Argentina. Reasignación de los recursos para el mejoramiento de la educación*. Washington D.C.: Serie de estudios del Banco Mundial sobre países; BANCO MUNDIAL (1986). *El Financiamiento de la Educación en los Países en Desarrollo*. Washington D.C.: Autor; BANCO MUNDIAL - GRUPO ESPECIAL SOBRE EDUCACIÓN SUPERIOR Y SOCIEDAD (2000). *El Buen Ejercicio del Poder*. En BANCO MUNDIAL, *La Educación Superior en los Países en Desarrollo: Peligros y Promesas*. Washington D.C.: Autor; TOMASEVSKI (2004). *El asalto a la educación*. Barcelona: Intermón Oxfam; CORAGGIO, & TORRES (1997). *La Educación según el Banco Mundial*. Buenos Aires: CEM/Miño y Dávila.

<sup>7</sup> El Consejo Federal de Cultura y Educación fue creado en el año 1972 por medio de la ley 19.682 y, según lo dispuesto en dicha norma, apuntaba a que la política educativa nacional se concrete coordinadamente «sobre la base de los estudios y experiencias realizadas dentro de los lineamientos fijados por el Plan Nacional de Desarrollo y Seguridad, ley 19.039, cuidando de respetar las necesidades de cada provincia». Por lo que no tenía atribuciones para fijar políticas educativas ya que sólo podía planificar y coordinar acciones entre el estado nacional y las jurisdicciones. Posteriormente, la ley 22.047 del año 1979 modificó su funcionamiento y organización. El gobierno del Presidente Alfonsín intentó sancionar leyes para modificar su funcionamiento durante los primeros años de su gestión pero no tuvo éxito en dicho emprendimiento ya que el Congreso Nacional no sancionó los dos proyectos de ley que se habían presentado en este sentido. Sin embargo, sí sancionó un decreto reglamentario de la ley 22.047 (Decreto PEN 934/84) que le dio su organización vigente hasta el año 2006 cuando fue reemplazado por el Consejo Federal de Educación.

Puede pensarse que la proclamada *devolución del poder a las instancias menores de gobierno*, por medio de la descentralización, respondió principalmente a una coyuntura ideológica e histórica más que a una necesidad pedagógica de modificación. Esta tendencia global a la reforma del Estado debería ser interpretada como una reestructuración de las relaciones entre sociedad y Estado. El caso de la descentralización educativa en la Argentina, efectuada durante la década de los noventa, se encuadró en este proceso de reestructuración del Estado y del cambio de su papel en la educación, por el que se pasó de uno *principal* a otro *subsidiario*.

La Ley Federal de Educación (LFE) constituyó un punto de inflexión en la historia de la educación argentina por muchos motivos pero uno es sobresaliente: la redefinición que establece en la relación entre el Estado Nacional y los estados federados en materia de prestación de la educación<sup>8</sup>. Esta modificación en la organización del gobierno originó diferentes situaciones entre los estados provinciales en lo que respecta a la implementación de dos aspectos novedosos (para la historia de la educación argentina) diseñados por esta LFE: la *estructura académica* y los *contenidos básicos comunes*. Las diferencias que se evidencian entre las jurisdicciones en la aplicación de estos dos aspectos de la reforma permiten pensar actualmente en la existencia de tantos sistemas educativos como jurisdicciones existan en la Argentina y quizás éste sea el aspecto más criticable de la ley 24.195<sup>9</sup>.

Estas situaciones de desigualdad en la nueva conformación del sistema educativo argentino se asientan precisamente sobre la *descentralización* de los servicios educativos. Así, la ley 24.049 constituyó la necesaria plataforma sobre la que se montó el diseño institucional de gobierno y administración previsto por la LFE. Consecuentemente, la

<sup>8</sup> La organización institucional del sistema educativo es el instrumento de la política educativa utilizado para estudiar la distribución de competencias y atribuciones entre el gobierno central y las jurisdicciones provinciales, en el caso de un estado federal, o bien entre la autoridad central y las instancias locales de gobierno, en los estados no federales. El estudio de la organización institucional de un sistema educativo apunta a analizar la estructura de las agencias de gobierno, los órganos intervinientes y las normas que los regulan, estudiándose así los núcleos de control de la educación. Por ello, las políticas de centralización y descentralización educativa forman parte del universo de análisis de la organización institucional.

<sup>9</sup> Para ilustrar esta gran diversidad, tomando el caso de los diseños curriculares de la Educación General Básica (EGB) derivados de los distintos grados y formas de aplicación de la LFE, por parte de las jurisdicciones, es posible encontrar muy distintas situaciones:

- Jurisdicciones con escuelas que ofrecen el 3° ciclo de la Educación General Básica (EGB3).
- Jurisdicciones con escuelas que ofrecen, en el mismo ámbito, EGB3 y Educación Polimodal.
- Jurisdicciones con escuelas que ofrecen Educación Polimodal.
- Jurisdicciones con escuelas que ofrecen, en el mismo ámbito, Educación Polimodal y Trayectos Técnico-Profesionales (TTP).
- Jurisdicciones con escuelas que ofrecen, en el mismo ámbito, EGB3, Educación Polimodal y TTP.
- Jurisdicciones con escuelas que ofrecen la educación secundaria tradicional.
- Jurisdicciones con escuelas que ofrecen educación técnica de nivel medio tradicional.

Para terminar de ilustrar esta situación de heterogeneidad educativa, podría mencionarse la situación de las distintas provincias respecto a la formación técnico-profesional (nivel medio) en particular:

- Provincias que están implementando masivamente los TTP.
- Provincias que están haciendo una implementación anticipada en algunas escuelas.
- Provincias que se propusieron la generalización de la implementación de los TTP a partir de 2001.
- Provincias que no están implementando los TTP ni se lo han propuesto en el año 2002.

Fuente: *Instituto Nacional de Educación Tecnológica*, 2001.

desatención de la población parecería ser el resultado de una modificación que, a través de la delegación de las responsabilidades en los estados provinciales sin las garantías económicas ni criterios pedagógicos, terminó por generar desigualdad educativa entre las jurisdicciones provinciales.

## **La descentralización educativa y sus efectos en la organización de la estructura académica del sistema educativo**

La LFE modificó radicalmente estructura académica hasta entonces vigente. La antigua estructura consistía en un *nivel inicial* no obligatorio (hasta los cinco años de edad), un único nivel obligatorio, constituido por *educación primaria* de siete años de duración, seguido por un nivel de *educación secundaria* de cinco o seis años de duración según las diferentes modalidades (bachiller, comercial, industrial, normal, entre otras). Finalmente estaba la *educación superior de estructura binaria*, conformada por *dos sectores: universitario y no universitario*, de duraciones variadas según las disciplinas. Todos los niveles dependían al momento de la sanción de la ley 24.195 de las jurisdicciones menos el superior universitario nacional, que seguía bajo la órbita del gobierno nacional. La LFE extendió la educación obligatoria a 10 años de duración y definió una nueva estructura académica que contemplaba:

- El nivel de *Educación Inicial*, cuyo último año pasaba a ser obligatorio; es decir, el correspondiente a los alumnos de cinco años de edad.
- La *Educación General Básica* (EGB) de nueve años de duración y de asistencia obligatoria, constituida por tres ciclos de tres años de duración cada uno (EGB1, EGB2 y EGB3). La definición de la duración de estos ciclos, los contenidos básicos de enseñanza respectivos y su ubicación surgió de los acuerdos federales dentro del CFCyE (posteriores a la sanción de esta ley) y de las disposiciones específicas que cada jurisdicción tomó para aplicar la ley.
- La *Educación Polimodal* no obligatoria. Con una duración que estuvo en torno a los tres años de duración (según las aplicaciones que realizaron los estados provinciales), sus regulaciones previeron cinco orientaciones generales posibles que surgieron de los acuerdos en el marco del CFCyE.
- La *Educación Superior* de duración variada según sean los sectores universitario y no universitario.

Lo problemático de este diseño fue su aplicación debido a que en función de la descentralización que garantizó la LFE las jurisdicciones provinciales eran las responsables

de implementar la nueva estructura académica y la política curricular. Los resultados fueron muy diversos debido a las capacidades financieras y de administración de cada jurisdicción. Por ende, las consecuencias fueron dramáticamente diferenciales, lo que promovió situaciones de alta desigualdad por las duraciones diferenciales de cada uno de los ciclos de la EGB. Finalmente, en 1995 el gobierno de Menem sancionó la ley 24.521, Ley de Educación Superior (LES), que avanzó sobre este esquema de gobierno del sistema educativo establecido por la LFE<sup>10</sup>. Paralelamente, y a través de los decretos reglamentarios de la LFE y de la LES así como también de los acuerdos logrados en el marco del CFCyE, se originó un contexto caracterizado, por un lado, por la implementación de una política de transferencia que no contempló una planificación educativa ni estudios de impacto como paso previo a la descentralización. Por otro lado, se generó un esquema de gobierno que recentralizó, en nuevas instancias u organismos ya existentes la definición de las políticas educativas, financieras y de evaluación de la calidad.

## La política educativa encarada a partir del año 2003

En la actualidad la cobertura del sistema educativo argentino comprende un total de 11.429.436 estudiantes en la educación formal, los cuales se hallan distribuidos de modo predominante en el sector público en todos los niveles pero sobre todo en el de educación básica (primaria) y superior no universitaria, tal como lo refleja el siguiente cuadro:

CUADRO I. Cobertura del sistema educativo argentino según niveles y sectores. Año 2005

Nivel	Matrícula		
	Valores absolutos	% Sector público	% Sector Privado
Inicial	1.324.529	69,7	30,3 %
Educación básica	5.324.807	77,7 %	22,3 %
Educación media	2.731.224	71 %	29 %
Educación superior no universitaria	509.134	53,9 %	46,1 %
Educación superior universitaria	1.539.742	83,5 %	16,5 %

Fuente: elaboración propia de acuerdo con datos de la Red Federal de Información (*Relevamiento Anual 2005. Ministerio de Educación, Ciencia y Tecnología*) y de la Secretaría de Políticas Universitarias (*Anuario 2005 de Estadística Universitaria. Ministerio de Educación, Ciencia y Tecnología. Año 2006*).

<sup>100</sup> Según la Ley de Educación Superior pueden mencionarse como órganos de gobierno, coordinación y consulta del sistema de educación superior los siguientes: el Ministerio de Educación de la Nación, los Consejos Regionales de Planificación de la Educación Superior, el Consejo Federal de Cultura y Educación, los Poderes Ejecutivos de las jurisdicciones; y para el caso de las universidades, en particular: el Consejo de Universidades, la Comisión Nacional de Evaluación y Acreditación Universitaria, el Consejo Interuniversitario Nacional y el Consejo de Rectores de Universidades Privadas.

El gobierno del Presidente Kirchner, al asumir el poder en mayo de 2003, inició una nueva reforma educativa. En este sentido, comenzó a sancionar legislación educativa desde la segunda mitad del año 2005. Estas leyes abrieron el panorama a la sanción de una nueva ley orgánica para la educación, lo cual ocurrió en diciembre de 2006. A continuación se analizan los aspectos sobresalientes de estas normas.

### **La ley 26.058 de Educación Técnico Profesional como intento ambiguo de subsanar uno de los vacíos generados por la Ley Federal de Educación**

El Congreso Nacional sancionó en septiembre de 2005 la ley 26.058. Es una norma que se caracteriza por la poca precisión de algunos de los conceptos vertidos en ella. Plantea definiciones generales sobre qué es la educación técnico profesional, en qué instituciones se provee este tipo de educación y a qué niveles del sistema educativo abarca. Sin embargo, sus definiciones son difíciles de comprender y cuando se las explicita se da cuenta de la falta de precisión sobre lo que efectivamente se debería enseñar en esta modalidad educativa. La ley plantea además un complejo sistema de organización institucional para la educación técnico profesional, lo cual se evidencia en la gran cantidad de organismos intervinientes en su definición y con competencias concurrentes pero también solapadas entre sí.

La ley apunta a regular la *educación técnico profesional* en el nivel medio y superior no universitario y lo que denomina *formación profesional* (art. 1) y se aplica al conjunto del sistema educativo nacional (art. 2) por lo que se invita a las jurisdicciones a adecuar su legislación en función de esta norma (art. 56). Declara que este tipo de educación constituye un derecho y la define como un servicio educativo profesionalizante que comprende la formación ética, ciudadana, humanístico general, científica, técnica y tecnológica (art. 3).

Los propósitos generales que presenta este tipo de educación presentan imprecisión conceptual. Así, se plantea que ella promueve: «el aprendizaje de capacidades, conocimientos, habilidades, destrezas, valores y actitudes relacionadas con desempeños profesionales y criterios de profesionalidad propios del contexto socio-productivo, que permitan conocer la realidad a partir de la reflexión sistemática sobre la práctica y la aplicación sistematizada de la teoría» (art. 4). Cabría esperar mayor rigor en la definición de conceptos tales como: «capacidades, conocimientos, habilidades, destrezas y actitudes», «desempeños profesionales» y «criterios de profesionalidad». ¿Por qué no denominar con precisión los aspectos que se regulan?

La ley intenta avanzar en lo que atañe a su cobertura institucional sobre la base de la fuerte desarticulación evidenciada en el sistema educativo nacional, sobre todo en este tipo de modalidad. Así, la ley sostiene que la educación técnico profesional «abarca, articula e integra los diversos tipos de instituciones y programas de educación para y en el trabajo» (art. 5), ya sean nacionales, jurisdiccionales, municipales, de nivel medio o superior no universitario y de formación profesional, de gestión estatal o privada (art. 9)<sup>11</sup>. Es posible distinguir tres tipos institucionales donde transcurre este tipo de educación (art. 9):

- Instituciones de educación técnico profesional de nivel medio.
- Instituciones de educación técnico profesional de nivel superior no universitario.
- Instituciones de formación profesional, constituidas por centros de formación profesional, escuelas de capacitación laboral, o equivalentes.

La ley distingue entre dos circuitos formativos:

- La *educación técnico profesional* que transcurriría en instituciones de nivel medio y superior no universitario y de la cual egresarían *técnicos medios* y *técnicos superiores* en áreas ocupacionales específicas (art. 7, inciso a). Es destacable también que se plantee que este tipo de circuito, en lo que se refiere al nivel superior no universitario, debería permitir «continuar con itinerarios profesionalizantes» (art. 12) aunque no se explicite qué se entiende por ello ni cómo se logrará ese propósito. Para el caso de la educación media se dispone que los planes de estudios deberán tener una duración mínima de seis años y que deben estructurarse según «los criterios organizativos adoptados por cada jurisdicción» (art. 24).
- La *formación profesional* que se daría en instituciones que no forman parte de la educación formal y que apuntaría «a preparar, actualizar y desarrollar las capacidades de las personas para el trabajo, cualquiera sea su situación educativa inicial (...) con inserción en el ámbito económico-productivo» (art. 8). Más adelante se dice que este circuito estaría constituido por «el conjunto de acciones cuyo propósito es la formación socio-laboral para y en el trabajo, dirigida tanto a la adquisición y mejora de las cualificaciones como a la recualificación de los trabajadores, y que permite compatibilizar la promoción social, profesional y

<sup>11</sup> Se evidencia el respeto por la terminología introducida en nuestra legislación en el año 1993 al considerarse la educación pública como «de gestión estatal».

personal con la productividad de la economía nacional, regional y local. También incluye la especialización y profundización de conocimientos y capacidades en los niveles superiores de la educación formal» (art. 17). Se trata además de un circuito que contempla «formas de ingreso y desarrollo» diferenciales a los requerimientos académicos propios de la educación formal (art. 18) a pesar que no se indica cuáles serían los criterios contemplados. No obstante, se plantea que estas instituciones y cursos «podrán ser reconocidos en la educación formal» (art. 20) pero no dice cómo.

En el capítulo IV del título III se puede apreciar la centralidad que adquiere el Consejo Federal de Cultura y Educación, que no deja de ser un consejo de ministros<sup>12</sup> en la regulación de esta educación técnico profesional. Este consejo es el que define la oferta académica específica (art. 21). Para ello debe considerar las «familias de profesiones» que elabore el Instituto Nacional de Educación Tecnológica (INET). En esta tarea el Consejo Federal debe establecer los criterios básicos y los parámetros mínimos referidos a: «*perfil profesional, alcance de los títulos y certificaciones y estructuras curriculares*, en lo relativo a la formación general, científico-tecnológica, técnica específica y prácticas profesionalizantes y a las *cargas horarias mínimas*» (art. 22). Dichos criterios serán los que encuadrarán los *procesos de homologación de títulos y certificaciones* de educación técnico profesional y servirán también para la estructuración de ofertas formativas o planes de estudio que pretendan para sí el reconocimiento de validez nacional por parte del Ministerio de Educación, Ciencia y Tecnología (art. 22). Asimismo, el Consejo Federal es el que define los «niveles de cualificación» como marco para la evaluación, reconocimiento y certificación de los saberes y capacidades adquiridos por los trabajadores en el ámbito laboral o por medio de modalidades educativas formales o no formales (art. 27)<sup>13</sup>.

Sin embargo, al igual que lo dispuesto por la LFE, en la ley 26.058 son las autoridades provinciales las que terminan diseñando los planes de estudios sobre la base de los criterios básicos y parámetros mínimos establecidos en el nivel federal y son ellas también las que establecen los requisitos de ingreso y las cargas horarias anuales y

<sup>12</sup> Este consejo fue reemplazado en diciembre de 2006 cuando se sancionó la Ley de Educación Nacional por el Consejo Federal de Educación. De todas formas su constitución y funciones son casi idénticas al anterior CFCyE. Está conformado por los responsables del área educativa de los poderes ejecutivos de los veinticuatro estados jurisdiccionales y tres representantes del Consejo de Universidades.

<sup>13</sup> Un tema muy serio está dado por la cuestión docente. En la ley 26.058 no se establece con precisión quiénes enseñarán en estos circuitos educativos. Se plantea en forma genérica que los docentes serán formados en el marco de programas federales de formación continua, los que a su vez serán concertados por el Ministerio de Educación Nacional con el Consejo Federal de Cultura y Educación, en lo que atañe a su implementación (art. 29).

totales. La aplicación de los contenidos básicos comunes derivados de la Ley Federal de Educación dejó como consecuencia de este esquema de definiciones curriculares un serio problema: la desarticulación horizontal del sistema que dio origen a una alta segmentación educativa. Ahora cabría esperar resultados similares ya que el contexto de empobrecimiento social no se ha modificado; nada hace pensar que los estados provinciales ejecutarán políticas diferentes a las que llevaron adelante cuando implementaron los diseños curriculares durante los noventa<sup>14</sup>.

Resulta sumamente complejo el proceso de reconocimiento oficial de las instituciones y la validación nacional de los estudios y títulos dentro de este circuito. En todos los casos el Ministerio Nacional y el Consejo Federal recentralizan las definiciones curriculares que encuadran las decisiones jurisdiccionales. Veamos cómo se organizan estas instancias:

- Registro Federal de Instituciones de Educación Técnico Profesional: integrado por las instituciones de Educación Técnico Profesional que incorporen las jurisdicciones (art. 32 y 34). El Instituto Nacional de Educación Tecnológica (INET) aparece como el responsable de gestionar este registro federal (art. 45, inciso e).
- Catálogo Nacional de Títulos y Certificaciones: es la nómina exclusiva y excluyente de los títulos y/o certificaciones profesionales y sus propuestas curriculares y está organizado en función de «familias y perfiles profesionales» (art. 32 y 36).
- Los títulos de técnicos medios y técnicos superiores no universitarios y las certificaciones de formación profesional podrán ser homologados en el orden nacional a partir de los criterios y estándares para el *proceso de homologación* acordados y definidos por el *Consejo Federal de Cultura y Educación*, los cuales deberán contemplar aspectos referidos a: perfil profesional y trayectorias formativas (art. 32 y 38).

Por su parte, la conformación de los órganos de administración y coordinación de la educación técnico profesional resulta también compleja (art. 41). Se evidencian las múltiples instancias y competencias solapadas:

<sup>14</sup> Es válido recordar que durante los años de implementación de la reforma educativa en la década pasada, como resultado de la distribución dispuesta por el régimen de coparticipación federal de impuestos, la mala administración y los ciclos económicos críticos vividos en la Argentina a partir de la crisis de 1995, varias jurisdicciones provinciales tuvieron severas crisis económicas que redundaron en la educación. Consiguientemente, se dejaron de pagar sueldos de los docentes, con los consecuentes paros y pérdidas de días de clases para los estudiantes de todos los niveles administrados por los gobiernos de los estados federados. Ello afectó la calidad de la educación en su conjunto y ciertamente limitó los alcances de la reforma educativa, al menos de sus objetivos declarados.

- El *Ministerio de Educación, Ciencia y Tecnología* aparece con centralidad en las definiciones políticas generales pero con necesidad de acordar las decisiones con el Consejo Federal (artículo 42°).
- El *Consejo Federal de Cultura y Educación* debe llevar adelante, entre otras, las siguientes funciones (art. 43):
  - Acordar los procedimientos para la creación y actualización de ofertas de educación técnico profesional.
  - Acordar los perfiles y las estructuras curriculares y el alcance de los títulos.
  - Acordar los criterios y parámetros de calidad hacia los cuales se orientarán las instituciones que integren el Registro Federal de Instituciones de Educación Técnico Profesional y los criterios y parámetros para la homologación de los títulos.
  - Acordar los procedimientos de gestión y distribución del Fondo Nacional para la Educación Técnico Profesional.
- Las *autoridades jurisdiccionales* son las que delinearán las características efectivas que adoptará esta modalidad en el país (art.44).
- El *INET* debe entre otras funciones (art. 45):
  - Llevar a cabo el relevamiento y sistematización de las familias profesionales, los perfiles profesionales y participar y asesorar en el diseño curricular de esta modalidad.
  - Administrar el Registro Federal de Instituciones de Educación Técnico Profesional, el Catálogo Nacional de Títulos y Certificaciones y llevar a cabo el proceso de Homologación de Títulos y Certificaciones.
- Se crea por esta ley también el *Consejo Nacional de Educación, Trabajo y Producción* como un órgano propositivo (art. 46). Entre sus funciones se destacan que apunta a conciliar intereses con los sectores productivos y actores sociales en materia de educación técnico profesional y promover la vinculación de la educación técnico profesional con el mundo laboral.
- También se crea por esta ley la *Comisión Federal de Educación Técnico Profesional* «con el propósito de garantizar los circuitos de consulta técnica para la formulación y el seguimiento de los programas federales orientados a la aplicación de la presente ley, en el marco de los acuerdos del Consejo Federal de Cultura y Educación» (art. 49); en este caso también el INET ejerce su coordinación (art. 50).

Como puede observarse, la ley conforma un complejo dispositivo de formación sin precisiones conceptuales para la gran cantidad de conceptos utilizados en la misma norma y con una alta estructura burocrática que se solapa en funciones.

## La ley 26.075 como intento de recomponer el financiamiento del sector

El 21 de diciembre de 2005 el Congreso Nacional sancionó la Ley de Financiamiento Educativo, la cual fue promulgada en enero de 2006 (Boletín Oficial 09/01/06). Es una norma que apunta a incrementar la inversión en educación, ciencia y tecnología por parte del estado nacional y las jurisdicciones de forma progresiva a partir del año 2006 hasta lograr una participación del 6% del producto bruto interno hacia el año 2010. Vale recordar que durante los años noventa la reforma educativa se ejecutó en un contexto de aplicación de medidas fiscales por las que el gobierno federal apuntó a un compulsivo «disciplinamiento fiscal», en función de las demandas que realizaron organismos internacionales de financiación<sup>15</sup>. Dentro de ese paquete de medidas fiscales se encontraba la transferencia a las jurisdicciones de los servicios educativos nacionales de nivel medio y superior no universitario, llevada a cabo luego de la sanción de la ley 24.049, entre los años 1992 y 1994<sup>16</sup>. De esta forma las jurisdicciones quedaron a cargo de una parte sustancial de la administración educativa, con capacidades diferenciales para llevar adelante esta tarea, en un contexto nacional de crisis económica y empobrecimiento generalizado.

Una de las disposiciones que se estableció en la LFE también implicaba un aumento gradual y sostenido de la inversión pública consolidada total en educación y se tomó como base el presupuesto destinado al sector en el año 1992 (art. 61 de la ley 24.195). El aumento suponía un incremento mínimo a razón de un 20% anual a partir del presupuesto del año 1993. Para ello, el Estado nacional y las jurisdicciones debían firmar un *Pacto Federal Educativo*, el que debía ser ratificado por el Congreso Nacional y

<sup>15</sup> En tal sentido deben ser comprendidas medidas tales como:

- el Pacto Fiscal firmado en el año 1992 (ley 24.130) por el cual el Estado nacional podía retener un 15% de los recursos de impuestos coparticipables para atender obligaciones previsionales a la vez que garantizó una suma fija para las provincias (\$725 millones mensuales) más allá de los niveles de recaudación;
- el Pacto Federal para el Empleo, la Producción y el Crecimiento del año 1993 que derogó impuestos provinciales e impuso a las provincias la obligación de reemplazar impuestos a los ingresos brutos por impuestos al consumo;
- la Ley de Compromiso Federal por el Crecimiento y la Disciplina Fiscal del año 1999 (ley 25.400) que dio lugar a una nueva suma fija de transferencia.

<sup>16</sup> Esta política muy perjudicial para la organización institucional del sistema daba cuenta de la voluntad del Estado nacional de desprenderse de la enorme masa de gasto público nacional para dejar que las jurisdicciones realizaran sus propios procesos de ajustes. De este modo se comenzaron a llevar adelante acciones que daban cuenta de las posiciones subsidiarias del Estado en materia de educación pública.

Se transfirieron instituciones de niveles que habían experimentado una fuerte expansión de sus matrículas respectivas, sin políticas de resguardo de la unidad y articulación, no sólo curricular, sino salarial. La descentralización de la negociación salarial profundizó el impacto heterogéneo de la transferencia en las provincias ya que puso en juego la capacidad de negociación de las jurisdicciones, así las últimas jurisdicciones en aceptar y firmar los convenios de transferencia fueron las que obtuvieron los mejores acuerdos por la presión que tenía el Estado nacional de finalizar el proceso, en función de sus compromisos internacionales que suponían, entre otras cosas, la reducción del peso de la administración central en el presupuesto nacional. Por otro lado influyó también, en la heterogeneidad de resultados obtenidos, las situaciones económicas y de recursos sociales diferenciados entre las jurisdicciones para sostener sistemas educativos ampliados que quedaron bajo su exclusiva responsabilidad de gestión.

por las legislaturas provinciales y apuntaría a la consecución de los objetivos de dicha ley (art. 63 de la ley 24.195). Este pacto se firmó en el año 1994. Lo interesante de todo este proceso es que lo dispuesto por la ley 24.195 y por el propio Pacto Federal fue luego derogado de hecho por las leyes de presupuesto nacional de los años subsiguientes, las cuales no cumplían con dichas disposiciones ante la necesidad de hacer frente a otras demandas, de sectores diferentes al educativo<sup>17</sup>.

En la ley 26.075 se dispone que el Ministerio de Educación de la Nación constituye la autoridad de aplicación de esta norma y por lo tanto debe acordar con las jurisdicciones dentro del Consejo Federal de Cultura y Educación las formas de implementación de las políticas tendentes al incremento de la inversión e incluso de los programas en los cuales se enmarcarán así como de las acciones que se desarrollarán para optimizar las capacidades de administración y evaluación y de la eficiencia del gasto sectorial (art. 11). En el art. 2 de la ley 26.075 se plantean varios objetivos a efectos de direccionar este proceso tendente al incremento de la inversión educativa, entre los que pueden destacarse los siguientes:

- Incluir en el nivel inicial al 100% de la población de cinco años de edad.
- Garantizar un mínimo de diez años de escolaridad obligatoria para todos los niños, y lograr que al menos un 30% de los alumnos de educación básica tengan acceso a escuelas de jornada completa.
- Avanzar en la universalización del «nivel medio/polimodal» y promover la reincorporación en él de los desertores.
- Erradicar el analfabetismo en todo el territorio nacional.

Para el cumplimiento de los objetivos se dispone que deberá crecer anualmente el gasto consolidado en educación, ciencia y tecnología del gobierno nacional. Se toma como base el año 2005 de acuerdo con los porcentajes consignados a partir de un cálculo, en el cual el gobierno nacional debe financiar con sus recursos los programas destinados a cumplir los objetivos especificados en los incisos referidos a la educación universitaria y las acciones de investigación del art. 2 de la ley en lo atinente a instituciones y organismos dependientes del Estado Nacional (art. 4).

En lo que respecta al resto de los objetivos explicitados en el art. 2, la ley estipula el incremento anual del gasto consolidado en educación, ciencia y tecnología de las

---

<sup>17</sup> El Pacto Federal Educativo dispuso el compromiso del gobierno nacional de invertir 3.000 millones de pesos en el lapso de cinco años, con una contraparte provincial del 20%. A esto se sumaba los aportes externos que financiaban diversos programas del Ministerio de Educación de la Nación y que estaban orientados a la implementación de la reforma educativa; la deuda generada por dichos fondos fue posteriormente asumida por el gobierno nacional.

jurisdicciones y también toma como base al año 2005, a través de un cálculo específico. En este caso, el incremento debe destinarse a la mejora de los salarios docentes, a la adecuación de las plantas orgánicas funcionales de las instituciones de modo de atender las necesidades de una matrícula incrementada y a la jerarquización de la carrera docente vía la capacitación (art. 5)<sup>18</sup>.

La ley también creó dentro del Ministerio de Educación de la Nación el *Programa Nacional de Compensación Salarial Docente* a efectos de subsanar las desigualdades salariales entre las provincias (art. 9). En este caso se delega al CFCyE la determinación de los criterios de asignación de los fondos para compensar las desigualdades evidenciadas entre las jurisdicciones a través del cálculo del porcentaje de los recursos determinados en el art. 4 de la ley que se destinen a este programa, así como su operatoria y los requisitos que deben cumplir las jurisdicciones para acceder a estos recursos (art. 9). Puede deducirse que el Consejo Federal de Cultura y Educación adquiere también centralidad en el control de la gestión de la aplicación de los recursos transferidos al ser considerado como el organismo «encargado de evaluar el funcionamiento del sistema de información física y financiera conforme a los clasificadores presupuestarios utilizados por la Ley 25.917 con el objeto de garantizar la homogeneidad de la información y el estricto cumplimiento de los compromisos entre las partes» (art. 16). Por último, la ley prorroga la vigencia del Fondo Nacional de Incentivo Docente por el plazo de cinco años a partir del uno de enero de 2004 (art. 19)<sup>19</sup>.

<sup>18</sup> Si bien se calcula el incremento de la inversión vía el Producto Interior Bruto (PIB) contemplado en la presentación del proyecto de Ley de Presupuesto de la Administración Nacional, se dispone que cuando no haya incremento en el PIB o cuando su variación «no genere el incremento en la recaudación exigible para alcanzar las metas financieras previstas, la meta anual deberá adecuarse proporcionalmente al incremento de la recaudación» y que se podrá acordar la redefinición de los plazos y condiciones de los compromisos asumidos por esta ley (art. 6). Para la obtención de estos fondos se dispone una asignación específica, por cinco años, de recursos coparticipables con la finalidad de garantizar la disponibilidad de los recursos previstos en el art. 5 de esta ley en los presupuestos de las provincias y de la Ciudad Autónoma de Buenos Aires. Asimismo, se establece que se afectarán, al incremento de la inversión, los recursos anuales coparticipables correspondientes a las jurisdicciones en el régimen de la ley 23.548 y sus modificatorias y complementarias (art. 7). Los compromisos de inversión sectorial anual por parte de las jurisdicciones provinciales deben respetar: «a) una participación del gasto en educación en el gasto público total no inferior a la verificada en el año 2005 y b) un gasto anual por alumno no inferior al verificado en el año 2005» (art.12).

<sup>19</sup> Este fondo había sido creado en el año 1999 por medio de la ley 25.919 y supuso un monto fijo que desde entonces se sumó al salario docente, con una suma adicional para el caso de nueve provincias con mayor urgencia por el deterioro salarial (Chaco, Chubut, Corrientes, Entre Ríos, Formosa, Jujuy, Misiones, Salta y Tucumán). La incorporación de este monto al presupuesto nacional implicó una modificación de la relación entre el estado nacional y las jurisdicciones en materia de financiamiento educativo respecto de la que había prevalecido en los años noventa por dos razones:

- El estado nacional comenzó a ofrecer una respuesta (más que una solución) a la demanda por una mejora del salario docente aunque a través de criterios de incentivos al trabajo.
- La financiación de este fondo dio lugar al abandono de otros sectores como infraestructura, capacitación docente y políticas compensatorias ya que las jurisdicciones perdieron las sumas que recibían para ellos de parte del estado nacional.

## La Ley de Educación Nacional: su análisis con especial énfasis en la nueva estructura académica propuesta

En la Argentina resulta necesario resolver los problemas derivados de la implementación de la estructura académica y de la política curricular luego de la reglamentación de la ley 24.195. La aplicación de la estructura académica diseñada en esa ley fue tan diferente en cada jurisdicción y dentro de cada una de ellas que puede pensarse en la pérdida del carácter sistémico de nuestro sistema nacional de educación. Esto es consecuencia de que perviven múltiples modalidades educativas gracias a la adaptación diferencial de dos de los ciclos y niveles más problemáticos que se diseñaron a partir de la ley 24.195: el 3° ciclo de la EGB (EGB3) y la educación polimodal. Es más, dos jurisdicciones no aplicaron la estructura académica diseñada por la LFE. El panorama se diferenció aún más en el año 2005 cuando la provincia de Buenos Aires desmanteló dicha estructura académica y creó dos niveles educativos nuevos: la «escuela primaria básica» (6 a 11 años de edad) y la «escuela secundaria básica» (12 a 14 años de edad) como unidades pedagógicas y organizativas propias y de asistencia obligatoria, la última como nivel previo a la educación polimodal<sup>20</sup>.

La reforma educativa del actual gobierno nacional, al igual que su antecesora, se caracteriza por una falta de diagnóstico integral de los problemas educativos y por la celeridad que se le dio al proceso de formulación de una nueva ley orgánica. Un aspecto muy importante que debe ser aquí considerado es el encuadre de este proyecto político educativo. En tal sentido, en la redefinición de una de las áreas centrales de la vida social de un país, como lo es la educación, cobra sentido siempre que se apunte a instalar un proyecto político hegemónico exitoso. En el mes de mayo de 2006 el Presidente de la Nación inició una convocatoria para un debate del conjunto de la sociedad vía la presentación de un documento de base y la sanción de un decreto que establecía plazos para la presentación de un nuevo proyecto de ley de educación nacional<sup>21</sup>. El calendario estipulaba plazos excesivamente cortos, tal como puede observarse en el siguiente cuadro:

---

<sup>20)</sup> Podría cuestionarse la legalidad de esta innovación en la medida en que la *escuela secundaria básica* (ESB) contradice la estructura académica de la ley 24.195 (vigente en el momento de la creación de la ESB) debido a que no respetaba la unidad pedagógica que dicha ley contemplaba para la Educación General Básica (EGB) y porque creó un nivel educativo que ella no preveía. Es más, se le dio entidad a la ESB a través del concurso de los cargos para su gestión directiva.

<sup>21)</sup> Presidencia de la Nación. *Ley de educación nacional. Hacia una educación de calidad para una sociedad más justa*. Documento para el debate. Buenos Aires: mayo de 2006.

---

**CUADRO II. Calendario fijado para el año 2006 por el gobierno nacional para la sanción de la Ley de Educación Nacional**


---

Actividad	Fecha/mes
Firma del decreto presidencial y lanzamiento del debate preparatorio de la nueva ley nacional	22 de Mayo
Fase de consultas a instituciones y especialistas	Mayo
Debate general, encuestas de opinión, jornadas en escuelas	Junio
Sistematización de los aportes y de la información disponible	Julio
Primera redacción del proyecto de ley	Agosto
Nueva fase de consultas a instituciones y especialistas	Septiembre
Presentación del anteproyecto de ley	16 de Septiembre
Redacción final y elevación del proyecto de ley al Congreso de la Nación	Octubre / Noviembre
Sanción de la nueva ley	Diciembre

Fuente: Ministerio de Educación, Ciencia y Tecnología

Al igual que en la década anterior la preparación del proyecto de ley se hizo en plazos muy cortos aunque con la diferencia de que en esta oportunidad su consenso fue mucho más exitoso al contarse con el aval de los gremios docentes en la redacción del proyecto de ley y por la legitimación que se dio al proceso de apertura de canales para que las personas, las instituciones y los *especialistas* pudieran acercar sus propuestas. Algunas jurisdicciones como la provincia de Buenos Aires incluso llevaron a cabo una consulta a los miembros de la comunidad educativa provincial: padres, docentes y alumnos. Dicha consulta fue realizada a través de una encuesta de opinión con categorías de respuestas discrecionales, que permitían obtener resultados que avalaran el proceso de consulta en su conjunto<sup>22</sup>. Asimismo, el documento de base para la convocatoria al debate presentaba precisamente «diez ejes de debate y líneas de acción» que tenían un alto nivel de generalidad sobre los aspectos centrales que cualquier ley de educación tendría que contemplar pero sin mayores precisiones sobre cómo serían regulados o no dichos aspectos. En suma, los términos de la convocatoria al debate y la forma en que se instrumentó, por su celeridad y canales efectivos de participación, resultan cuestionables, siendo prácticamente nula su validez.

---

<sup>(22)</sup> En agosto de 2006, la Dirección General de Cultura y Educación de la provincia de Buenos Aires hizo circular un documento titulado «Consulta 2006. Hacia una nueva ley de educación nacional», que estaba destinado a padres y alumnos y debía ser administrado por las autoridades de las instituciones, es decir, debían realizar jornadas de reflexión, encuestar a los padres y alumnos y procesar los datos para su envío a la dirección general. La encuesta tenía tres preguntas: en la primera se preguntaba si la escuela secundaria debía ser obligatoria; en la segunda si ella debía formar (y daba varios ítems de respuestas posibles) para: a) construir identidad individual e identidad colectiva; b) seguir estudiando; c) ingresar en el mundo laboral; d) desempeñarse como ciudadano activo. En la tercera pregunta se indagaba si «para promover una buena educación es necesario»: a) respeto por los derechos de los alumnos; b) condiciones laborales dignas para los docentes; c) la capacitación y actualización de los docentes; d) la participación de los padres; e) valoración de la cultura del trabajo y del esfuerzo; f) la valoración de las áreas de la expresión artística y de la educación física. Como puede advertirse las categorías de respuestas resultan sumamente cuestionables ya que no resultan mutuamente excluyentes, orientan la respuesta y los encuestados no poseen información ni posibilidad para elaborar respuestas alternativas. Más allá que desde un punto de vista valorativo es casi impensable no acordar con semejante nivel de vaguedad contemplado por cada uno de estos ítems.

El proyecto de ley que el gobierno nacional presentó el 16 de septiembre tenía 139 artículos, distribuidos en 12 títulos. Lo más sobresaliente de la propuesta resultó la modificación de la estructura académica que había establecido la LFE a tan solo 13 años de su creación. Son escasos los antecedentes internacionales que den cuenta de cambios de tal envergadura por parte de un sistema educativo nacional en tan breve lapso de tiempo, sobre todo porque los cambios en las políticas educativas tienen implicaciones en el largo plazo histórico. Es más, es deseable implementar cambios graduales que no perturben la unidad del sistema en su conjunto. Este proyecto fue sancionado como *ley 26.206, Ley de Educación Nacional (LEN)* con 145 artículos distribuidos en 12 títulos<sup>23</sup>.

Más allá de todos los aspectos que regula esta ley y que requieren ser tratados con detenimiento, aquí se analizan como un análisis preliminar las disposiciones relativas a la estructura académica, la formación docente y al gobierno y administración del sistema<sup>24</sup>. En el articulado de la ley sobresalen aspectos tales como: la modificación en la estructura académica, las políticas de formación docente y la centralidad que tiene el (por esta norma creado) *Consejo Federal de Educación*, que reemplaza al Consejo Federal de Cultura y Educación que había sido creado en 1972. De hecho el último título de esta ley establece la derogación de la Ley Federal de Educación y de las leyes que regulaban el funcionamiento del CFCyE. La composición de este Consejo Federal resulta compleja debido a que se contempla la existencia de miembros plenos con voz y voto mientras que también hay representantes sectoriales sin capacidad de voto (artículo 117°, inciso a). Se destaca también que las resoluciones que sancione este consejo serán vinculantes para las jurisdicciones pero nuevamente se induce a la confusión al establecerse que el carácter de «cumplimiento obligatorio» de dichas resoluciones quedará sujeto a la disposición de la Asamblea Federal del consejo (su máximo órgano de gobierno) y de acuerdo con la reglamentación que emita el consejo (art.

<sup>23</sup> El proyecto de la Ley de Educación Nacional ingresó al Senado proveniente del Poder Ejecutivo Nacional y se publicó en el Diario de Asuntos Entrados N° 190 con fecha del 16 de noviembre de 2006. Sus firmantes fueron el Presidente de la Nación Néstor Kirchner, el Jefe de Gabinete de Ministros Alberto Fernández y el Ministro de Educación, Ciencia y Tecnología Daniel Filmus. En el Senado fue remitido a las comisiones de Educación, Ciencia y Tecnología, de Población y Desarrollo Humano, y de Presupuesto y Hacienda. El día martes 5 de diciembre de 2006 se aprobó el proyecto con modificaciones al original en una reunión conjunta de las comisiones a las que había sido remitido y un día después (6 de diciembre) tuvo media sesión al ser aprobada en el recinto de la Cámara de Senadores. El mismo miércoles 6 de diciembre de 2006 el proyecto pasó a la Cámara de Diputados donde fue enviado a las comisiones de Educación y de Presupuesto y Hacienda, las que en una reunión conjunta celebrada el 7 de diciembre, lo aprobaron sin modificaciones con tres dictámenes en minoría. El 14 de diciembre de 2006 la Cámara de Diputados del Congreso Nacional aprobó el proyecto, publicado como ley en el Boletín Oficial el día 28/12/2006.

<sup>24</sup> La ley contiene regulaciones importantes sobre: el principio de inclusión educativa para la educación especial; la educación rural; la educación intercultural bilingüe con formación docente específica (aunque vagamente definida); la educación en contextos de privación de la libertad; la educación domiciliaria y hospitalaria; disposiciones curriculares sobre la inclusión de contenidos relativos a la soberanía sobre las Islas Malvinas y al proceso histórico y político relativo a un período de terrorismo de estado entre los años 1976 y 1983; entre otros.

118). Es decir, la ley no especifica en qué oportunidades o circunstancias las resoluciones del Consejo Federal se volverán obligatorias para las jurisdicciones y ello queda a criterio discrecional de sus miembros, reunidos en la asamblea y en función de las correlaciones de fuerzas que poseen cada uno, más que en función de las políticas o temas a regular.

En lo que a la *estructura académica* se refiere, la ley elimina la EGB y la Educación Polimodal y dispone el *retorno* a las denominaciones preexistentes: *Educación Primaria* y *Educación Secundaria* con la curiosa distinción que prevé la coexistencia de *dos posibles estructuras académicas en función de su duración* (art. 134):

- Una opción de seis años de duración para el nivel primario y otros seis años para el nivel secundario;
- Otra posibilidad que contempla siete años de duración para la educación primaria y cinco años para la secundaria, tal como era la estructura vigente históricamente hasta la sanción de la ley 24.195 y que continua intacta en la ciudad de Buenos Aires, por tomar un caso de no aplicación de dicha norma.

La ley contempla la obligatoriedad de la educación secundaria (art. 29), junto con el último año de la educación inicial (art. 18). Por lo que la obligatoriedad de la escolaridad se extendería a 13 años en total. De esta forma la nueva estructura del sistema educativo tendría cuatro niveles educativos: inicial, primario, secundario y superior y ocho «modalidades» educativas: técnico profesional; artística; especial; permanente de jóvenes y adultos; intercultural bilingüe; rural; y educación en contextos de privación de la libertad; domiciliaria y hospitalaria (art. 17)<sup>25</sup>.

La aplicación de la nueva estructura se torna problemática en el nivel de educación secundaria, lo que da cuenta de la gravedad de la situación en materia de enseñanza debido a la adopción diferenciada por parte de las jurisdicciones de la EGB3 y de la educación polimodal. Allí es donde incluso los escasos y parciales análisis que desarrolló el Ministerio de Educación de la Nación dan cuenta de la diversidad de duración de los ciclos y niveles, agravado por la diferencia de diseños curriculares en lo

<sup>25</sup> Se evidencia aquí una confusión entre modalidades educativas (en cuanto a contenidos) y regímenes de enseñanza que adquiere la educación formal común. Rigurosamente definidas, las modalidades son las divisiones que se establecen en el interior de un nivel del sistema educativo basadas, específicamente, en un campo de conocimiento. De esta manera encontramos, por ejemplo, en el nivel medio orientaciones que reúnen a ciertas disciplinas desarrollando propuestas curriculares que, después de una orientación general, ofrece trayectos paralelos de formación específica. En consecuencia, la *educación rural* o la *educación en contextos de privación de la libertad*, por mencionar dos casos, constituirían en rigor *regímenes especiales* de la educación común más que modalidades diferenciadas de ella.

que atañe no sólo a la definición de las áreas curriculares sino también a cantidad de horas semanales y anuales destinadas a su enseñanza en cada jurisdicción<sup>26</sup>.

La LEN lejos de apuntar a equiparar, al menos en el largo plazo, la situación de diferenciación horizontal detectada, contempla la posibilidad de existencia de dos «opciones» de estructuras académicas, para las cuales no especifica las formas de articulación y absorción de las situaciones existentes en cada jurisdicción. Al contrario, delega al Ministerio Nacional y al Consejo Federal de Educación la atribución de especificar los criterios para la implementación de estas políticas relativas a la estructura académica. Con ello se asemeja al esquema previsto por la ley 24.195, la cual no definía algunas cuestiones relativas a estos mismos aspectos, cuando podía haberlos regulado y evitar así las consecuencias que conllevó su aplicación diferencial por parte de los estados jurisdiccionales. En este caso, la ley prevé un plazo de seis años, para que «a través de acuerdos entre el Ministerio de Educación, Ciencia y Tecnología y el Consejo Federal de Educación, se defina la ubicación del séptimo año de escolaridad. El Ministerio de Educación, Ciencia y Tecnología y el Consejo Federal de Educación acordarán los criterios de unificación que, respetando las condiciones de las distintas jurisdicciones, aseguren los mecanismos necesarios de equivalencia y certificación de los estudios, movilidad de los/as alumnos/as y derechos adquiridos por los/as docentes» (art. 134). Sin embargo, no se define la forma y los procesos de homologación curricular y de títulos que se deberán respetar para mantener la unidad del sistema. Más aún, tampoco se dispone explícitamente dónde se deberá ubicar el séptimo grado (¿formará parte de la educación primaria o estará en la secundaria?), ni tampoco si todos los estados federados tendrán que ubicarlo en el mismo nivel educativo: ¿algunos podrán ubicarlo en la educación primaria y otros en la secundaria?

Dos aspectos importantes relativos también a la estructura académica aquí diseñada tienen que ver con la ratificación de normas ya existentes. Por un lado, el capítulo destinado a la educación técnico profesional cuenta con un solo artículo en el que se ratifica la vaga organización delineada para esta modalidad por medio de la ley 26.056 (art. 38), anteriormente analizada. Por otro lado, el capítulo destinado a la educación superior ratifica la vigencia de la muy cuestionada ley 24.521 de Educación Superior (art. 34), que fue objeto de fuertes críticas por parte de la comunidad universitaria democrática

---

<sup>26)</sup> Véase al respecto, entre otros, los siguientes trabajos: MINISTERIO DE EDUCACIÓN DE LA NACIÓN (2000). *Estado de la situación curricular en las provincias*. Buenos Aires: Subsecretaría de Educación Básica – Programa de gestión curricular y capacitación; MINISTERIO DE EDUCACIÓN DE LA NACIÓN (2002). *Estado del arte sobre la implementación del Tercer ciclo de la Educación General Básica*. Buenos Aires: Dirección Nacional de Información y Evaluación de la Calidad Educativa – Unidad de Investigaciones Educativas. CIPPEC (2004). *Mirada comparada de la reforma educativa en las provincias: Fichas por provincia*. Buenos Aires: Área de política educativa del CIPPEC.

durante el proceso de sanción y luego durante su traumática implementación. Entendemos que la vigencia de la ley 24.521 constituye un aspecto fuertemente cuestionable debido a que esta norma reglamentarista supone un modelo competitivo entre las instituciones de este nivel, que lo fragmenta a la par que centraliza acciones en el complejo y oneroso sistema de gobierno y coordinación que allí se prevé.

Por otro lado, a pesar de estas disposiciones la Ley de Educación Nacional contempla la posibilidad de existencia de instituciones de educación superior no universitarias bajo dependencia nacional en el mismo capítulo referido a la educación superior. También crea el *Instituto Nacional de Formación Docente* dentro del Ministerio de Educación de la Nación como órgano consultivo y propositivo tendiente a homogeneizar criterios curriculares para la formación docente inicial y continua (art. 76). Sin embargo, la ley no indica la naturaleza jurídica de este instituto, ni los orígenes de los fondos que manejará, ni establece tampoco controles y requisitos para la administración de esos fondos. Se limita a enumerar las funciones que tendrá a cargo (como órgano propositivo, no resolutivo) pero no incluye entre ellas a la directa *formación de docentes*, aspecto clave en cualquier reforma educativa.

## Conclusiones

La política educativa del gobierno de Kirchner se ha caracterizado por plantear y llevar adelante un cambio normativo estructural, con el propósito de derogar las normas que habían modificado el sistema educativo en los años noventa. Con esta apuesta el gobierno también intenta diferenciarse de aquella cuestionable reforma educativa, encarada como parte del proceso de reforma del Estado y que supuso la desfinanciación del sector educativo. Así es que se presentan los principios y propósitos que se enuncian y los derechos que se reconocen a los integrantes de la comunidad educativa en las leyes analizadas en este artículo. Resulta cuestionable la forma en que se implementarán las profusas declamaciones que se presentan en estas leyes, sobre todo en la LEN, ya que el contenido normativo propiamente dicho resulta escaso. Ello se debe a que no se definen con precisión muchas de las regulaciones que deberían explicitarse para hacer efectivos los principios, propósitos y derechos mencionados. Las palabras no producen cambios por sí solas. Así, no por incluir artículos que comiencen con los verbos «garantizar» y «asegurar», la ley logrará que se cumplan. De hecho la Ley

Federal de Educación también planteaba una educación de calidad, ampliaba el período de obligatoriedad de los estudios y sostenía derechos para los docentes. Es más, la LFE dio lugar a un Pacto Federal Educativo para incrementar la financiación del sector pero ello no redundó en asignaciones dignas para los trabajadores docentes ni tampoco en niveles de capacitación y formación que permitieran llevar a cabo la transformación declarada en los años noventa.

Por último y al igual que la LFE, esta nueva ley marco para el sistema educativo delega en el nivel federal de gobierno del sistema, constituido por el Consejo Federal de Educación, la definición de los aspectos críticos de la política educativa. Con ello se vuelve a apelar al consenso aunque la ley destaca que puede haber decisiones de cumplimiento obligatorio para todos los estados jurisdiccionales. No obstante, como ya se señaló, no se precisa en qué circunstancias ello ocurrirá. ¿Cuándo un acuerdo del CFE será de cumplimiento obligatorio? ¿Qué porcentaje de acuerdo será requerido? ¿Mayoría simple? ¿Unanimidad? Vuelve a ser cuestionable también que sea el CFE el que tenga la potestad para las definiciones centrales de la política educativa nacional cuando esto es una atribución del Congreso de la Nación, de acuerdo con el art. 75, inciso 19 de la Constitución Nacional.

En consecuencia, varias preguntas se derivan de lo descrito en los párrafos anteriores: ¿cuándo comenzará a regir la nueva estructura? ¿Cómo se capacitará a los docentes en ejercicio para este tipo de estructura académica? ¿Qué se hará con los cargos docentes concursados para ciclos y niveles que dejarán de existir? ¿Cómo se adecuará la planta funcional de las instituciones en función de la nueva estructura? ¿Cómo se definirán los diseños curriculares para los nuevos niveles? ¿Se adaptarán los vigentes para los cuales los docentes habían sido capacitados previamente y el Estado había gastado muchos recursos? ¿Qué posibilidades hay de que no se acreciente la diferenciación horizontal en cada nivel cuando se prevé desde la misma ley la coexistencia de dos estructuras académicas de diferente duración? Valdría recordar, para esto último en especial, que en la educación el orden de los factores altera el producto. Ante estas pocas preguntas puede pensarse que las leyes analizadas en este trabajo, sobre todo la LEN, no resultan operativas. No se regula con claridad los aspectos referidos a las definiciones curriculares, la conformación precisa de la estructura académica y la formación de los docentes. Se utilizan muchas palabras y no se las define con precisión y se deja abierta la posibilidad a que esta vaguedad de lugar a problemas de interpretación.

Si no se contemplan estos aspectos problemáticos, si no se consideran las condiciones efectivas en las que se desarrolla la enseñanza en cada contexto institucional y

regional, es poco probable que se alcancen algunos de los objetivos declarados por la nueva reforma educativa argentina. Pero también resulta muy importante aprender las lecciones que dejó la aplicación de la ley 24.195. Más allá que el contexto de sanción de esta norma no es comparable con el vigente en el año 1993, debido al aval que tiene la actual reforma por parte de las cúpulas gremiales de los docentes, sí parecen comparables algunas características que la nueva ley: la falta de previsión efectiva de las eventuales consecuencias que tendría una modificación tan profunda de la estructura académica (en el marco de niveles de gobierno con competencias concurrentes). No hay similitud en cuanto al contenido de la estructura académica que esta nueva ley sustenta ya que se deroga la creada por la ley del año 1993. Sin embargo, parece ser idéntica la forma que tendrá su aplicación en el conjunto del sistema educativo nacional. Ello en sí mismo constituye un problema muy serio cuyas consecuencias no parecen ser previstas adecuadamente.

**Dirección de contacto:** Guillermo Ruiz. Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET). Universidad de Buenos Aires, Argentina. E-mail: [gruiz@psi.uba.ar](mailto:gruiz@psi.uba.ar)



# Los procesos cognitivos y meta-cognitivos en la composición escrita del alumnado de etnia gitana con desventaja socioeducativa

## Cognitive and metacognitive processes in writing composition by gypsy pupils with socio-educational disadvantages

Antonio García Guzmán

*Universidad de Granada. Facultad de Ciencias de la Educación. Departamento de Didáctica y Organización Escolar.  
Granada, España*

### Resumen

En este artículo se presenta una investigación cuyo objetivo es analizar la expresión escrita del alumnado de etnia gitana con desventaja socioeducativa, escolarizado en Educación Primaria. Concretamente esta investigación, siguiendo el modelo socio-cognitivo de Flower y Hayes (1981) y Hayes (1996), pretende desvelar los procesos cognitivos y meta-cognitivos que estos sujetos llevan a cabo en la redacción de un texto. Se trata de una investigación de carácter cualitativo, concebida como un «estudio de caso», compuesto por dieciocho alumnos de etnia gitana con desventaja socioeducativa en el que se utiliza tanto la «entrevista cognitiva» como la «corrección de errores en textos» para la obtención de datos y se aplica el «análisis de contenido», para la interpretación de los mismos. Esta investigación ha permitido identificar las operaciones que estos alumnos muestran en el proceso de producción de un texto (procesos cognitivos de planificación, transcripción y estructuración, así como los procesos meta-cognitivos), el modo en el que los alumnos ejecutan tales operaciones y, sus dificultades para expresarse por escrito.

Los resultados de la investigación han sido interpretados, haciendo una lectura transversal (es decir, teniendo en cuenta los datos obtenidos en diversas fuentes o en distintos apartados de la misma fuente), así como estableciendo diferencias y semejanzas, en función de las

características diferenciales de los alumnos, en el interior del grupo. Por último, se marcan una serie de directrices didácticas dirigidas a profesores y formadores con el objetivo de mejorar y/o subsanar aquellas dificultades que han sido descritas en dichos procesos y que interfieren en el desarrollo de la competencia escritora y se realiza una prospectiva de esta investigación.

*Palabras clave:* composición escrita, alumnos de etnia gitana, procesos cognitivos, dificultades en la escritura, educación primaria, desventaja socioeducativa.

### **Abstract**

This paper is aimed at analyzing the writing expression of gypsy pupils with socio-educational disadvantages attending primary school in Granada (Spain). Specifically, this research relies on Flower and Hayes' (1981) and Hayes' (1996) socio-cognitive framework to unveil the processes that gypsy children undergo when they write a text. This qualitative research entails a 'case study' of eighteen gypsy pupils with socio-educational disadvantages. Both 'cognitive interviews' and 'mistake-spotting tasks' are used to gather data, while 'content analysis' is used to interpret them. This form of data interpretation permitted identifying the operations that these pupils follow in the process of writing a text (e.g. planning, transcribing and structuring, as well as the metacognitive processes), the way in which pupils perform such operations and the difficulties they encounter when producing a text.

The results of this research have been interpreted by cross-reading sets of data (namely, taking into account data obtained from several sources or from different sections within the same source) and identifying differences and similarities across sets of data, bearing in mind differential characteristics of the research subject.

Finally, teachers and trainers are provided with some teaching guidelines aimed at overcoming and/or correcting the difficulties experienced by gypsy children (as described above) and which interfere with the development of pupils' writing skills. Further research implications are discussed.

*Key Words:* writing composition, gypsy pupils, cognitive processes, writing difficulties, primary education, socio-educational disadvantages.

## **Introducción**

El objetivo de esta investigación ha sido analizar los procesos cognitivos y meta-cognitivos que los alumnos de etnia gitana con desventaja socioeducativa utilizan en la composición escrita. Los resultados obtenidos forman parte de una investigación más amplia, cuyo objetivo ha sido analizar la expresión escrita del alumnado de etnia

gitana, escolarizado en Educación Primaria. La importancia de dicha investigación ha sido avalada con la concesión de una beca de Formación del Profesorado Universitario (FPU), para el período 2004-07, auspiciada por el Ministerio de Educación y Ciencia, y en la actualidad por la Universidad de Granada.

Los hechos que han motivado la realización de este estudio son varios: en primer lugar, el interés científico de esta línea de investigación en la composición escrita, y la investigación en esta área con sujetos procedentes de diferentes *minorías étnicas o grupos sociales minoritarios* y con diversos problemas socioeducativos, como es el caso del alumnado de etnia gitana, entre los que cabe destacar: elevado fracaso escolar, absentismo escolar, bajo rendimiento..., aún habiéndose producido un gran avance en los últimos años, principalmente en su escolarización (FSG, 2002).

En segundo lugar, lo poco que se ha estudiado la expresión escrita con alumnos de etnia gitana, a pesar su incidencia en el desarrollo cognitivo (Vygotsky, 1978). Este aprendizaje, además, se encuentra totalmente conectado con factores socio-culturales (Sperling y Friedman, 2001). De otra parte, el aprendizaje de la expresión escrita, junto al aprendizaje matemático, son los contenidos más importantes del currículum, a la vez que esenciales en el ser humano, de ahí que se les considere como aprendizajes instrumentales.

## Antecedentes y fundamentación teórica

La investigación sobre la escritura está en desventaja, en comparación con otras áreas del currículum. «En España se ha investigado poco sobre procesos cognitivos en la expresión escrita» (Salvador, 2005, p. 11), aunque la investigación sobre características textuales sea algo más abundante. Sin embargo, en los últimos años, la investigación en la escritura ha sido reconocida como un importante y valioso campo de estudio, como así lo muestran los dos manuales de investigación más relevantes publicados sobre este tema, el *Handbook of Research on Teaching* (Richardson, 2001) y el *Handbook of Writing Research* (MacArthur, Graham y Fitzgerald, 2006).

En la investigación sobre la expresión escrita se han adoptado diversos criterios que derivan en tres enfoques principales, los cuales son complementarios y no excluyentes, pues cada uno analiza y explica una parte del complejo proceso de la escritura y no su totalidad: 1) el enfoque de proceso se centra en el sujeto (el escritor), es decir, en el proceso de producción textual, y es de carácter predominantemente

psicológico; 2) el enfoque centrado en el producto se ocupa del texto y es de carácter lingüístico; 3) el enfoque contextual (o ecológico), de carácter didáctico, centra la atención en el contexto en que se produce el aprendizaje. Estos enfoques, en efecto, pueden y deben ser complementarios (Berninger y Whitaker, 1993). Esta clasificación se corresponde, en parte, con una perspectiva temporal. Así, en la década de los setenta, el centro de atención eran los aspectos formales de la escritura; en los ochenta, emerge la orientación cognitiva; y en la década de los noventa, se produce un notable incremento de programas de enseñanza, centrados en el enfoque contextual.

A pesar de que las investigaciones en español sobre los procesos cognitivos son escasas, ya se cuenta con algunas iniciativas relevantes, como las llevadas a cabo por el Grupo de Investigación ED.INVEST de la Universidad de Granada (véase Salvador, 2005). Existen, además, en nuestro contexto, otros autores también con dilatada experiencia en este campo de estudio (Camps y Ribas, 2000; Castelló, 1999; Cuetos, Sánchez y Ramos, 1996; García Sánchez y Fidalgo, 2003; González Seijas, 2000; Teberosky, 1992; Tolchinsky, 1993; Vilá, 1998).

Las causas que justifican la ausencia o escasez de investigaciones con estos sujetos en esta materia curricular son variadas. Entre ellas, la consideración de que la cultura gitana es de tradición oral (Poveda, Cano y Pomares-Valera, 2005; Smith, 1997). En segundo término, cuando la alfabetización se ha tomado como objeto de estudio en comunidades gitanas, es porque ésta se ha considerado como un índice de aculturación hacia la sociedad mayoritaria y, por tanto, la adquisición de la lecto-escritura han cobrado relevancia en sus usos oficiales y vernáculos (San Román, 1990).

En otras áreas, como la lectura, aparecen algunas investigaciones, aunque también son escasas (Oriente, 1979). No obstante, son más frecuentes las investigaciones sobre el lenguaje, en general, y más concretamente sobre la funcionalidad lingüística de estos sujetos (Bueno, 1993). En relación con éstas, el estudio de la expresión escrita (adquisición y dificultades) se haya menospreciado.

Otras investigaciones (Poveda, Cano y Palomares-Valera, 2005) han examinado la escritura espontánea de los alumnos de etnia gitana prestando atención a la forma del texto, el contenido del escrito y el contexto social en el que el texto tiene aplicación. Los datos mostraron cómo, cuando el contexto y el contenido de la escritura se vinculan con su experiencia personal, estos sujetos son capaces de construir textos complejos. Por otro lado, Hernández (1996), siguiendo con la línea de investigación centrada en el enfoque de producto, evaluó la expresión escrita del alumnado de etnia gitana, comparándolos con otros niños no gitanos, a través de pruebas estandarizadas (TECI). En la mayoría de las categorías, el grupo de la comunidad gitana mostró resultados inferiores.

En referencia a los procesos cognitivos, no se han desarrollado estudios que analicen cómo son las operaciones que se producen en el niño de etnia gitana cuando está escribiendo, aún menos su desarrollo y dificultades.

## El enfoque cognitivo de la composición escrita

Existen diversos modelos cognitivos que explican el proceso de composición escrita pero, sin duda, el más conocido y elaborado, y el que mejor explica este proceso en su conjunto y, sobre todo, el que mejor pone de relieve los componentes cognitivos del proceso de producción es el modelo de Flower y Hayes (1981), ampliado por Bereiter y Scardamalia (1987) y De Beaugrande (1984) y, progresivamente modificado y perfeccionado por Hayes (1996). Según este último, la escritura es el resultado de una adecuada combinación de factores cognitivos, afectivos, sociales y físicos.

En el modelo cognitivo prototípico de Flower y Hayes (1981) se distinguían tres unidades:

- La memoria a largo plazo
- El contexto de la tarea
- Los procesos cognitivos implicados en la escritura:
  - la planificación (decidir qué decir y la forma de decirlo)
  - la transcripción (proceso que convierte la planificación realizada en texto escrito);
  - la revisión o relectura (subdividido en dos subprocesos: la evaluación del resultado y la revisión y corrección de lo producido).

Los procesos cognitivos descritos, a su vez, están controlados o regulados por otros procesos de más alto nivel: los procesos meta-cognitivos, referidos al conocimiento y control que tiene un sujeto sobre los procesos cognitivos de planificación, transcripción y revisión (Flower y Hayes, 1981; Hayes, 1996).

En definitiva, los procesos meta-cognitivos hacen referencia a si el sujeto es consciente de la conducta, del conocimiento y de las emociones y de si es capaz de controlarlos, adecuándolos a las exigencias de la situación y de la tarea. La función de supervisión, que caracteriza a estos procesos, permite la interacción y la recursividad de los procesos ejecutivos (planificación, transcripción y revisión) y establece reglas de prioridad, de secuencia y de interrupción entre los procesos (Salvador, 2005). A continuación, se describen los contenidos o referentes de los procesos meta-cognitivos y las dificultades que pueden surgir en su desarrollo: conocimiento del proceso,

de la estructura textual, de sus capacidades y de los procesos de autorregulación.

Especialmente relevantes son el conjunto de investigaciones de Bereiter y Scardamalia (1987) que, a partir de investigaciones de tipo experimental, identifican procesos de composición escrita diferentes, que formulan en los modelos de «decir el conocimiento» y de «transformar el conocimiento». Según este modelo los escritores, dependiendo de su dominio y experiencia en el arte de escribir, utilizarán un determinado modelo. Los alumnos inexpertos utilizarán modelo básico «decir el conocimiento» y los más expertos el modelo avanzado y completo «transformar el conocimiento». No obstante, también se da la posibilidad de combinar ambos modelos o estadios para responder a las exigencias de las tareas de escritura.

Desde esta propuesta, se han sucedido otras, realizadas por éstos y otros autores (Bereiter y Scardamalia, 1987; Hayes, 1996) en los que se perfilan y describen las operaciones implicadas en la escritura pero se acepta la clasificación anterior. Podría aseverarse que las operaciones de planificación, transcripción y revisión son ampliamente aceptadas, incluso en nuestro panorama, aunque su inclusión ha sido demorada en relación con otros contextos.

## Diseño y metodología de investigación

En la metodología de investigación sobre los procesos cognitivos, coincidiendo con el cambio en el objeto de estudio, del proceso al producto, y caracterizado por un enfoque más cualitativo, la técnica de investigación más utilizada en el enfoque de proceso ha sido la de «hacer pensar en voz alta» a los sujetos, mientras realizan la composición o inmediatamente después (Emig, 1971). La entrevista clínica puede ser una alternativa al «pensamiento en voz alta». En ella se puede utilizar un cuestionario, para ayudar a reflexionar a los sujetos que tienen dificultades para verbalizar su pensamiento.

Del mismo modo, la obtención de datos también ha experimentado cambios sustanciales. Éstos han de derivarse de las técnicas utilizadas en el campo de la Lingüística y la Psicolingüística, especialmente en el enfoque de proceso, en el que se ha ubicado la presente investigación. Las técnicas utilizadas por el enfoque procesual para la obtención de datos sobre la composición escrita son las siguientes (Salvador, 2000):

- vía directa o análisis de los procesos;

- vía indirecta o análisis del producto, es decir, el texto escrito. Ésta consiste en analizar las características estructurales y formales de los discursos escritos por el alumno con la intención de descubrir, a partir de ellas, los procesos cognitivos que ha seguido el alumno.

Ambas técnicas son complementarias, pues analizan dos aspectos (proceso y producto) de un mismo fenómeno: la escritura. La investigación que se presenta a continuación está centrada en el análisis de los procesos, sin entrar en el análisis del producto, es decir, aunque en este trabajo de investigación también se les pidió a los alumnos que escribieran un texto narrativo de libre elección, para posteriormente analizarlo a través de las técnicas de la gramática textual, en este artículo sólo se presentan los resultados obtenidos mediante el análisis de los protocolos de las entrevistas y de los dos textos que se les da al sujeto para que identifique y corrija determinados errores léxicos, ortográficos, de puntuación, sintácticos y estructurales, con el objetivo de contrastar los resultados en el proceso de revisión.

La investigación cualitativa puede considerarse como una metodología apropiada para el estudio y conocimiento de la composición escrita; muchos de los avances significativos producidos en los últimos años en este campo provienen de estudios cualitativos (Schultz, 2006), aspecto que justifica la aplicación de dicho método en esta investigación. Esta investigación cualitativa se ha concebido como un «estudio de caso múltiple», cuyo referente han sido 18 alumnos y alumnas de etnia gitana con desventaja socioeducativa, escolarizados en Educación Primaria. Este estudio de caso se caracteriza por su particularización y no por su generalización formal, ya que se estudian los procesos cognitivos y meta-cognitivos en la escritura de un grupo de sujetos en situación de desventaja socioeducativa y pertenecientes a la comunidad gitana y se llega a conocerlos en profundidad, sin necesidad de establecer comparaciones para ver en qué se diferencian de otros alumnos (Stake, 2005). Sin embargo, los resultados aquí obtenidos ayudarán a conocer más detalladamente la naturaleza de estos procesos, por lo que se puede hablar de una cierta «transferabilidad» o «extrapolación del conocimiento adquirido» (Gobo, 2004). Es por ello que este estudio de caso es al mismo tiempo un *estudio de caso múltiple* (estudio de un grupo de sujetos en particular) e *instrumental* (Stake, 2005), ya que, además, esta investigación contribuye a desarrollar la teoría psicolingüística sobre el aprendizaje de la escritura.

Del mismo modo, es necesario resaltar que, como cualquier estudio de caso, esta investigación está orientada al estudio del caso (el caso se entiende como un sistema acotado); evita el reduccionismo y el elementalismo; es relativamente no comparativo y busca comprender su objeto, más que comprender en qué se diferencia de otros (Stake, 2005).

## Problema de investigación y objetivos (acotación e intereses)

Dado el carácter exploratorio-descriptivo de este estudio, las hipótesis, derivadas de diversos modelos teóricos (lingüísticos, psico-lingüísticos y socio-cognitivos), equivalen a interrogantes fundamentales («erotemas») sobre distintas operaciones psicológicas producidas en estos sujetos, implicadas en la construcción de textos, y sobre las características del texto construido.

En general, esta investigación pretende conocer y desvelar:

- ¿Qué procesos cognitivos y meta-cognitivos utilizan en la expresión escrita estos alumnos de etnia gitana en situación de desventaja educativa, y cómo interactúan en el acto escritor?
- ¿Qué dificultades encuentran en la expresión escrita dichos alumnos?

A partir del modelo teórico descrito (Flower y Hayes, 1981; Hayes, 1996) y de las aportaciones de otras investigaciones (Berninger y Whitaker, 1993; Graves 1975; Salvador, 2005) se construyó un sistema de categorías correspondientes a las dimensiones psicolingüísticas analizadas (véase Cuadro I).

CUADRO I. Sistema de categorías utilizado para el análisis de los protocolos de las entrevistas.

«PROCESOS COGNITIVOS EN LA COMPOSICIÓN ESCRITA»		
Planificación (P)	Transcripción (T)	Revisión (R)
Procesos de pensamiento del sujeto sobre el contenido, la organización y los objetivos de la composición escrita.	Procesos del sujeto para desarrollar la estructura y la forma del texto.	Procesos del sujeto para evaluar el texto escrito, de acuerdo con una planificación previa.
P1-Génesis de las ideas.	T1- Ordenación sintáctica.	R1-Adecuación de la forma y/o el contenido a lo planificado.
P2-Auditorio.	T2-Riqueza de vocabulario.	R2-Estructura y léxico de la oración.
P3-Objetivos/ finalidades / intención.	T3- Selección léxica.	R3-Puntuación, ortografía.
P4-Selección de ideas.	T4-Adecuación de las palabras a las ideas.	R4-Caligrafía.
P5- Secuenciación de ideas.		R5- Revisión por otros.
P6-Fuente de las ideas.		R6- Revisión por sí mismo.
P7-Registro de ideas.		
E1-Organización textual: tipo de texto.		
E2-Organización general del texto.		
PROCESOS META-COGNITIVOS		
AR1- Conocimiento y control de la planificación.		AR5- Disposición ante la escritura y sus dificultades.
AR2- Conocimiento y control de la transcripción.		AR6- Conocimiento del escrito bien hecho.
AR3- Conocimiento y control de la revisión.		AR7- Conocimiento y control general del acto de la escritura.
AR4- Conocimiento y control de la estructuración.		

Fuente: Salvador y García, 2005: p. 68, modificado.

## Estrategias y técnicas para la obtención de datos

Para la obtención de datos *sobre los aspectos estudiados*, se siguieron procedimientos diferenciados, en función de las dimensiones abordadas en la investigación. En todas ellas, la observación directa en el proceso de recogida de datos ha sido de gran relevancia.

- **Entrevista** semiestructurada de tipo cognitivo. En ésta, el investigador interroga al sujeto, inmediatamente después de construir un texto narrativo, para conocer las estrategias que sigue el escritor en la ejecución de los procesos cognitivos, así como el control y conocimiento que tiene sobre éstos (metacognición). Se partió de un *cuestionario-guía*, compuesto por 69 preguntas (Salvador, 2000, p. 66-71), progresivamente mejorado (véase Salvador y García, 2005, p. 62-66) con el objetivo de dar algunas pistas a los alumnos sobre los procesos cognitivos activados durante la construcción de un texto. A continuación, se muestran algunas de las cuestiones que componían dicho cuestionario (Cuadro II).

CUADRO II. Sistema de categorías utilizado para el análisis de los protocolos de las entrevistas

PROCESOS COGNITIVOS	
<b>PLANIFICACIÓN</b>	3) Antes de escribir un texto o una redacción, ¿piensas en la persona que lo va a leer? 7) ¿Usas alguna fórmula o truco para recoger y ordenar las ideas que se te ocurren y no olvidartas mientras escribes? 13) ¿Cómo consigues escribir un texto o una redacción tú solo? ¿Qué es lo primero que haces? ¿y después...?
<b>TRANSCRIPCIÓN</b>	21) Cuando tienes dificultades para encontrar la palabra adecuada, ¿qué haces: buscar otra parecida, escribirla sin pensar, usar alguna regla? 22) ¿Las palabras de tus textos o redacciones las vas escribiendo siguiendo algún orden?
<b>REVISIÓN</b>	34) Cuando has escrito el texto, ¿se lo das a leer a un compañero o a otra persona para que te lo revise? 36) Cuando estás revisando un texto que acabas de escribir, ¿te fijas si a las oraciones que has escrito les faltan palabras?
PROCESOS META-COGNITIVOS	
<b>ESTRATEGIAS DE AUTO-CONTROL</b>	49) Mientras estás escribiendo, ¿te planteas qué tipo de texto vas a escribir y lo que tienes que hacer para que esté bien escrito? 55) ¿Te controlas mientras escribes? Si estás nervioso, cansado o disgustado ¿procuras superar la situación y concentrarte en lo que estás escribiendo? 58) Cuando te atascas ¿consigues seguir adelante? ¿Cómo lo haces?
<b>CONOCIMIENTO SOBRE LA ESCRITURA</b>	61) Suponte que te piden que hagas de profesor hoy en tu clase y que uno de los alumnos te pregunta: ¿Qué es escribir bien? ¿Qué le dirías a ese alumno? 62) ¿Crees que hay compañeros que escriben mejor que tú? ¿Cómo piensas tú que lo hacen los compañeros que escriben bien? 67) ¿Crees que después de escribir se debe revisar el texto? ¿Para qué?

Fuente: Salvador y García, 2005: p. 68, modificado.

- **Análisis de errores.** En esta técnica el alumno trata de detectar errores en textos escritos por otros. A este efecto, en el contexto de la entrevista se le presentan al sujeto dos textos narrativos muy breves, escritos deliberadamente con errores de diverso tipo (de puntuación, ortográficos, sintácticos, morfológicos y estructurales) (véase Salvador y García, 2005, p. 66), para que los señale y enumere aquellos que podrían afectar a la comprensión del texto. El

entrevistador leyó el texto en voz alta, mientras el alumno lo leía en voz baja. Al alumno se le dió la siguiente consigna: *El texto lo ha escrito un compañero tuyo. Si otro compañero tuyo tuviera que leer este texto, podría tener dificultad para comprenderlo, porque tiene algunos errores. Intenta descubrir cuáles son los errores que pueden hacer difícil la comprensión del texto y corrígelos, redactando nuevamente el texto, de forma correcta.*

### **Proceso de recogida de datos**

Para hacer que la obtención de datos resultara lo más «normalizada» posible, previa a la aceptación del alumno para participar en la investigación, se le explicó el proceso e interés de la misma, así como los materiales utilizados, las pruebas y la duración aproximada. Igualmente, las pruebas se aplicaron en el aula de apoyo o en lugares conocidos por el alumno, y se procuró crear un ambiente distendido y una situación agradable. Durante el proceso de obtención de datos, se anotó cualquier incidencia, gestos, sentimientos y conductas que pudieran ser relevantes para el análisis posterior de los datos.

### **Estrategias y técnicas para el análisis de datos**

En el análisis de datos se utilizaron diversos procedimientos. En primer lugar, los protocolos de las entrevistas, debidamente transcritas, en los que se recogían las reflexiones del escritor, se sometieron al *análisis de contenido* (Bardin, 1986; Krippendorff, 1997), que consistió, básicamente, en clasificar la información verbal contenida en una serie de documentos, de acuerdo con un sistema de categorías. El análisis de los protocolos permitió hacer patentes los diversos procesos utilizados en la escritura y las relaciones entre ellos.

El *sistema de categorías conceptuales establecido* derivó del modelo teórico de Flower y Hayes (1981) y Hayes (1996), descrito anteriormente. Se definieron operativamente varias categorías y subcategorías, que se corresponden con los procesos ejecutivos básicos en la composición escrita. Del mismo modo, se asignaron códigos a las categorías y a las operaciones o subcategorías.

Tras el análisis de los protocolos de las entrevistas, se agruparon los contenidos (o enunciados) cuyo código era similar y luego se agruparon los códigos de una misma categoría en las meta-categorías establecidas: planificación, transcripción, revisión y meta-cognición (véase Cuadro I). A continuación, se obtuvieron frecuencias y porcentajes correspondientes en cada una de las categorías y meta-categorías. Por último, se abordó la interpretación global de los datos obtenidos, tanto los de tipo numérico como los de tipo verbal.

Para facilitar la evaluación e interpretación de los datos obtenidos en la corrección de textos, realizada por el alumno, se elaboró un sistema de categorías y se procedió del mismo modo que con los datos obtenidos en las entrevistas: agrupación de códigos asignados a cada categoría y obtención de frecuencias. Las categorías definidas fueron: puntuación (P), ortografía (O), sintaxis (S), morfología (M), léxico (L), y estructura textual (E).

En el análisis se tuvo en cuenta tanto el aspecto cuantitativo como el cualitativo. En cuanto al primero, se obtuvo la cantidad de aciertos en la corrección de cada uno de los tipos de errores. Con estos datos, se establecieron comparaciones entre los distintos tipos de errores corregidos y entre los sujetos, en función de sus características diferenciales.

De otra parte, con respecto al análisis cualitativo, se distinguió entre la identificación de errores, si el alumno es capaz de detectar el error, y la corrección de los mismos, es decir, si, una vez detectado, el alumno es capaz de corregirlo. Para ello, se realizó un contraste entre los textos revisados y modificados por el alumno y los textos correctos. En definitiva, se trataba de evaluar la habilidad mostrada por el alumno para corregir los errores, lo cual, a su vez, permite inferir su conocimiento sobre el texto bien construido.

Tanto el cuestionario-guía de la entrevista como el sistema de categorías para el análisis de contenido se validaron mediante el procedimiento del *juicio de expertos*. El cuestionario fue valorado por distintos investigadores, miembros del grupo ED.INVEST, de la Universidad de Granada. Estos jueces determinaron la validez de contenido del instrumento, de acuerdo con los criterios de relevancia de los ítems y su referencia (o representatividad) a los diversos procesos en la composición escrita.

De igual modo, el sistema de categorías fue sometido al juicio de los mismos expertos, para comprobar si dicho sistema reunía las siguientes características:

- exhaustividad: las categorías permiten clasificar la totalidad de la información recogida en las entrevistas;
- exclusión mutua: cada unidad de registro no puede incluirse en más de una categoría;
- homogeneidad: las categorías están definidas, de acuerdo con un mismo principio;
- pertinencia: las categorías se adaptan a la información contenida en las entrevistas y al modelo teórico;
- productividad: el conjunto de las categorías proporcionan resultados valiosos, tanto en cuanto a las inferencias como en cuanto a nuevas hipótesis.

Para asegurar la validez y credibilidad de los resultados tres investigadores expertos realizaron el análisis de contenido de los protocolos de las entrevistas. El procedimiento para calcular la fiabilidad de los análisis fue la *triangulación de codificadores independientes* (Fox, 1981). Este procedimiento consiste en calcular el porcentaje de veces que varios codificadores independientes coinciden en la asignación de códigos a las unidades de registro, establecidas en un mismo material. En este caso concreto, el porcentaje de acuerdo fue del 96%. Para calcular este porcentaje se contó el número de veces que los codificadores coincidían en la asignación de un código a una unidad de registro y se convirtió ese número en porcentaje, en relación con el total de unidades de registro codificadas.

La fase final en el proceso de análisis de datos fue de especial importancia, por cuanto se trataba de interpretarlos de forma global. A este efecto, los datos fueron interpretados haciendo una lectura transversal (es decir, teniendo en cuenta los datos obtenidos en diversas fuentes o en distintos apartados de la misma fuente).

## Sujetos de la investigación

Los *casos seleccionados* corresponden a 18 alumnos de etnia gitana (10 chicos y 8 chicas) escolarizados en Educación Primaria y cuyo nivel socio-económico y cultural familiar es medio-bajo, ya que los progenitores de quince de estos sujetos tienen sólo estudios primarios, dos tienen estudios medios y por último, un padre y una madre carecen de estudios. Del mismo modo, sólo dos madres tienen trabajo, y los padres, excepto uno de ellos que no tiene, trabajan principalmente en la venta ambulante, trabajos temporales y en la construcción.

Se ha partido de alumnos que, siendo de etnia gitana, presenten un cierto desarrollo de los procesos cognitivos, en general, y de la escritura, en particular. De esta manera, se han seleccionado a alumnos de entre 10 y 13 años. La medida de edad es de 11,3 años de edad.

En lo que respecta al nivel educativo de los alumnos, 17 sujetos se encuentran escolarizados en el tercer ciclo de Educación Primaria (trece de ellos en 6º de Primaria y cuatro en 5º), por último, un sujeto en el segundo ciclo de la Primaria (4º de Primaria). De éstos, siete presentan dificultades de aprendizaje y la mitad (nueve sujetos) recibe clases de refuerzo en aula de apoyo.

Los sujetos proceden de dos colegios de carácter concertado de la zona norte de Granada. Del colegio «A» doce sujetos y del colegio «B» seis, de los que cinco tienen dificultades de aprendizaje.

## Análisis y discusión de resultados

En este apartado se contemplan los resultados obtenidos en atención a las variables psicolingüísticas estudiadas y se establecen algunas comparaciones entre los resultados obtenidos con estos alumnos y los obtenidos con otros alumnos de Educación Primaria, en los que el diseño de investigación ha sido similar (Arroyo y Salvador, 2005; Salvador y García, en prensa).

En términos generales, todos los sujetos planifican las ideas y piensan lo que van a poner, antes de escribir, aunque con diferentes matices y utilizando diferentes contextos para generarlas. Sin embargo, sólo seis sujetos clasifican u ordenan las ideas que finalmente escriben en el texto, aunque también varían en la forma de hacerlo. Estos resultados difieren a los obtenidos con otros alumnos (Arroyo y Salvador, 2005, p. 367), donde la mayoría de los alumnos sí *revelaron una claridad y organización muy próxima a lo que debe ser el dominio del proceso escritor*.

Un elevado número de sujetos (14) tiene en cuenta la persona a la que se dirige el texto: *Yo quiero la letra muy muy... colocada, bien escrita para que la entienda la maestra* (sujeto 11/p. 2/ P2); *Como mi hermana es más pequeña habría que hacerle cosas de niños chicos ¿no? y a ti po pa más grandes* (sujeto 16/p. 10/ P2-AR3).

Los objetivos que se plantean al escribir son que el texto les salga bien y les sea reconocido: *¿Yo? Pues que se presente bien la letra, y que lo entienda la maestra, los alumnos de mi clase y que me digan que lo he hecho muy bien* (sujeto 4/p. 7/P3); *Que salga en la tele, en los periódicos, que la maestra los mande a otros colegios* (sujeto 6/p. 7/P3). Sin embargo, es también elevado el número de sujetos (14) que escribe principalmente por obligación, lo cual denota una falta de motivación hacia la escritura.

Los alumnos del estudio, en general, carecen de estrategias para elaborar un determinado tipo de texto. Así, en la planificación, sólo cuatro sujetos utilizan algún esquema o cuadro para organizar las ideas (sujetos 5, 7, 9, 14) y, dos sólo algunas veces (sujetos 6 y 8). Aspecto que se hace aún más visible en el análisis de los procesos meta-cognitivos, ya que un amplio número de sujetos (once) no tiene ningún tipo de estrategia para que el texto les salga bien desde el principio. Los otros sujetos, aunque verbalmente afirman utilizar algún tipo de estrategia en determinadas ocasiones (sujetos 2, 6, 11 y 12) o siempre (sujetos 15, 17 y 18) más que utilizarla, ponen en marcha procesos básicos y necesarios al redactar un texto, como pensar antes de escribir (sujetos 6, 11 y 18), o procesos mecánicos como *sacarle punta al lápiz* (sujeto 2/p. 5/AR6). Sólo es destacable la estrategia que utiliza el sujeto 17: *Po... escribir en la hoja aparte y si me salen palabras y luego ya las sé, po las pongo bien* (sujeto 17/p.

7/AR6). Sin embargo, en el estudio ya citado (Arroyo y Salvador, 2005) los alumnos sí utilizan la estrategia del registro de ideas.

En la organización general del texto, aunque un amplio número de sujetos (trece) afirma ordenar las ideas y/o palabras de forma distinta en función del tipo de texto, los resultados obtenidos en la corrección de textos difieren, ya que no detectan errores sobre la estructura del texto ni corrigen aspectos relacionados con la misma. Los sujetos que corrigen algún aspecto en esta categoría lo hacen incorporando referentes en las frases.

Del mismo modo, los resultados obtenidos en los procesos meta-cognitivos revelan un escaso conocimiento sobre la estructuración de los textos y la ordenación sintáctica de las frases y párrafos dentro de los diferentes tipos de texto. En algunos casos, sobrealoran su capacidad de estructuración o, realmente, no tienen claro la estructuración que los diferentes textos conllevan (quince sujetos), o simplemente creen no saber cómo se construye un texto cuando les ha salido mal. Estos resultados fueron también obtenidos en una investigación de características similares (García, 2003).

Cuando se les pidió que describiesen el proceso que llevarían a cabo para elaborar un texto descriptivo, argumentativo, comparativo, narrativo o periodístico (una noticia), los tipos de texto en los que encontraban menor dificultad fueron en los textos de carácter narrativo, y descriptivo, pese a que su estructura era simple y en la mayoría de los casos carecían de un orden lógico. Por otro lado, tenían un desconocimiento total del texto argumentativo, así como dificultades para organizar la información en un texto periodístico. Resultados que confirman el escaso conocimiento que poseen de la estructura general de un texto.

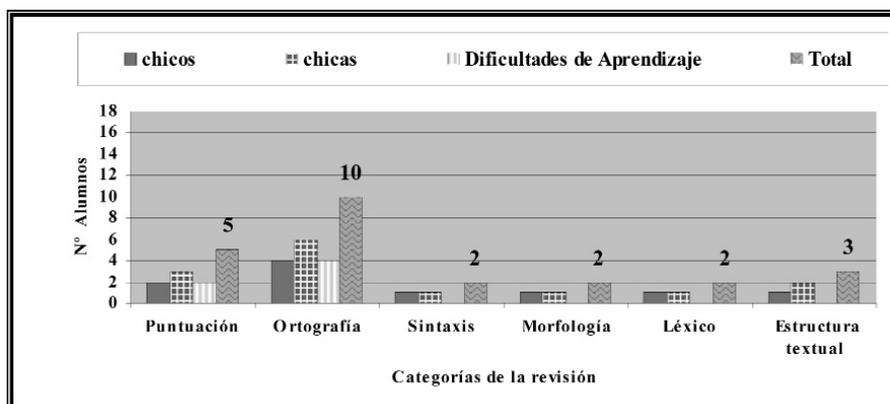
Casi la totalidad de los sujetos tiene algún tipo de dificultad para encontrar las palabras adecuadas al redactar un texto. Así, tanto la observación en el proceso de evaluación (tiempo invertido en la producción del texto y en las cuestiones formuladas al investigador) como los resultados en la corrección de errores en los textos corroboran las dificultades que estos alumnos presentan en la selección léxica o su escaso vocabulario. Por ejemplo, en este último, sólo dos sujetos corrigen algún aspecto del léxico en los textos. Este aspecto se corrobora en otra investigación, llevada a cabo con alumnos de etnia gitana, donde se obtuvo que el vocabulario de los alumnos de etnia gitana era considerablemente inferior al de los alumnos no gitanos (Hernández, 1996).

En general, los sujetos son conscientes de la importancia que tiene revisar sus textos. Además, los que creen necesario revisar sus textos coinciden en afirmar que el objetivo de esta revisión es ver si está bien o mal lo que han escrito y/o detectar los errores. Sin embargo, como se ha ido comprobando, sólo son capaces detectar y/o

corregir aspectos referidos a la forma y no al contenido o a la estructura general del texto. De este modo, se limitan principalmente a buscar palabras que estén mal escritas, con errores ortográficos y/o de puntuación. Así cuando cambian aspectos del texto se centran más en aspectos formales que en aspectos sintácticos o estructurales: *Los puntos, las comas...* (Sujeto 6/p. 7/R6-R3); *pues si tengo una palabra que empieza por la «h» y la tengo mal, pues la corrijo y la pongo bien* (sujeto 3/p. 4/R1). En comparación con otro estudio, realizado con alumnos de Educación Primaria (Salvador y García, en prensa), los alumnos de dicha investigación tenían en cuenta, además de aspectos ortográficos y de puntuación, determinados aspectos léxico-semánticos y morfológicos. Incluso en otros estudios, algunos alumnos, menores de 10 años, han sido capaces de revisar su texto en profundidad, haciendo cambios en la estructura que afectan al significado (Dix, 2006).

Estos resultados son semejantes a los obtenidos en la corrección de textos hecha por los alumnos, ya que en las categorías de ortografía y puntuación se producen mayor número de correcciones, aunque no son muchas y, en algunos casos, en lugar de disminuir el número de errores incluidos los han aumentado. Los alumnos no son capaces de detectar errores sintácticos, morfológicos, léxicos o de estructura textual. Tan sólo dos sujetos corrigen aspectos relacionados con la sintaxis y la morfología de las palabras (sujetos 5 y 8); también, sólo dos (sujetos 8 y 11) corrigen algún aspecto léxico y, tres sujetos, la estructura textual (sujetos 5, 8 y 9), incorporando el referente omitido en determinadas frases («abuela» o «cuento»), pero ninguno corrige otros aspectos de la estructura textual, como ordenar las frases o los párrafos dentro del texto (véase Gráfico I).

GRÁFICO I. Resultados sobre la corrección de textos



Fuente: elaboración propia

Los sujetos tienen dificultades notables en la fase de detección de errores más que en la de corrección, ya que los que detectan los errores normalmente los corrigen bien. Sin embargo, la mayoría no encuentra lo que está mal y, en ocasiones, intentan corregir sin éxito algún aspecto, como es la puntuación, pero, más que disminuir el número de errores, los aumenta.

Estos resultados son propios de escritores inexpertos, ya que este proceso de revisión es el que más se descuida (Salvador, 2000). No obstante, tampoco este aspecto está muy desarrollado, ya que la proporción de errores detectados y corregidos es mínima, en comparación con los que fueron incluidos en los textos.

Por otro lado, aunque la mayoría acepta y ve positivo que le digan que han de corregir algún aspecto del texto, a seis sujetos (sujetos 2, 4, 11, 13, 15 y 18) no les gusta que les digan que su texto está mal: *Mu mal maestro, ya lo tengo becho y me dicen que...* (sujeto 15/ p. 7/AR5). *pues mal...estoy muy nerviosa y las leo las palabras y las ves si están mal las palabras...porque si la maestra te pone muy bien yo ya me siento, yo no sé, más feliz, porque a veces tengo las cosas mal hechas y me pone un regular o un mal* (sujeto 11/p. 3/AR5). Aquí, además, se puede ver cómo la actitud del alumno hacia la escritura se relaciona con el método de enseñanza.

Todos los sujetos son conscientes de lo que sienten mientras escriben y de que su estado de ánimo o actitud hacia la escritura influirá en el resultado final del texto. Un aspecto que podemos destacar es que hay un importante número de sujetos que, en ocasiones, cuando escriben se encuentran nerviosos. Además, son conscientes de que, cuando están nerviosos, cansados o enfadados, la calidad de sus textos es inferior: *A veces me sale peor, porque me pongo muy nerviosa* (sujeto 11/p. 6/AR5).

Gran parte de los sujetos afirma concentrarse en lo que quieren escribir, aunque, en ocasiones, cuando están escribiendo se atascan y, consecuentemente, algunos se agobian, dejan de escribir o no saben qué hacer (sujetos 4, 7, 8, 9, 15 y 18).

El conocimiento de lo que es escribir bien y los procesos que implica parece no estar muy claro. La mayoría tiene dificultad para expresar qué es escribir bien o por qué hay compañeros que escriben mejor que ellos. De ahí que un elevado número de sujetos entienda que escribir bien sea no tener faltas de ortografía o tener una buena caligrafía. Justificación dada también para explicar por qué hay compañeros que escriben mejor que ellos: además de reconocer que se esfuerzan más, atienden más y están menos nerviosos cuando escriben. Por dicho motivo, estos alumnos creen que para mejorar sus escritos han de poner más atención, no estar nerviosos, esforzarse más (leer y escribir en su casa) y, por tanto, pensar más y buscar información o ideas en diferentes fuentes de datos: *Leer muchos textos y pensar y escribir*

*mejor...ir escribiendo mucho y que cada vez salga mejor la letra* (sujeto 8/ p. 4/AR6); *Escribir más en mi casa* (sujeto 10/ p. 6/AR6).

Es de destacar, además, la importancia que algunos alumnos conceden al maestro como la clave en la mejora de su competencia escritora: Pues que la maestra insista en mí para que escriba mejor, o que me mande (sujeto 4/p. 5/AR6).

Por último, atribuyen la baja calidad de su texto a:

- las faltas de ortografía,
- mala construcción de frases o inadecuada selección léxica,
- poca atención o esfuerzo al redactar el texto.

## Conclusiones y proyección didáctica

De esta investigación se derivan aportaciones al conocimiento sobre los procesos cognitivos implicados en la escritura de alumnos de etnia gitana con desventaja socioeducativa:

- planifican las ideas y piensan lo que van a poner, antes de escribir, aunque esta operación aparece principalmente en textos narrativos;
- tienen una buena génesis de ideas;
- tienen en cuenta a quién va dirigido el texto (auditorio);
- son conscientes de todo lo que sienten cuando escriben (si están nerviosos, cansados...), y que su estado de ánimo influirá en la calidad final del texto;
- son conscientes de la importancia de la revisión en el proceso escritor.

Por otro lado, esta investigación ha permitido detectar, igualmente, las dificultades que estos alumnos encuentran al realizar estas operaciones:

- la mayoría escribe por obligación. Sin embargo, les gusta que se les reconozca el trabajo realizado;
- tienen dificultades para clasificar y ordenar las ideas en el texto;
- presentan un escaso conocimiento sobre la estructuración general del texto, así como de la ordenación sintáctica de oraciones y párrafos;

- poseen un vocabulario pobre;
- se observan dificultades para detectar y/o corregir errores de puntuación, ortografía y, más concretamente, en el contenido o estructuración general del texto;
- la revisión se centra en aspectos formales (puntuación, ortografía y caligrafía);
- muestran escaso conocimiento sobre lo que es escribir bien y lo que esto conlleva, aunque son conscientes de que la mejora de su competencia escritora depende de la atención y esfuerzo que pongan al aprender y, en cierto modo, del profesor;
- evidencian gran desconocimiento de estrategias para mejorar la composición de sus textos.

En consonancia con los resultados obtenidos en esta investigación, se derivan implicaciones educativas para los profesionales de la educación, para los formadores y asesores del profesorado, para los diseñadores de materiales curriculares y para las instituciones educativas en la orientación hacia una política de desarrollo lingüístico para la atención a grupos sociales con desventajas socioeducativas:

- Partir de una metodología de enseñanza en la que se *primen los logros académicos*, por muy pequeños que sean y sean positivamente valorados, tanto por la familia como por el profesorado.
- Crear situaciones donde el alumno aprecie la escritura como algo divertido, agradable, que necesita relajación y concentración y, en definitiva, que se sienta atraído por ella. Aspecto que es olvidado en muchos de los programas de instrucción de la escritura (De Caso y García Sánchez, 2006) y que, como se ha visto en el estudio, resulta esencial para estos alumnos.
- Trabajar diversos tipos de textos en el aula (noticias, textos argumentativos, comparativos), con el propósito de que los alumnos conozcan su estructura, y valoren la importancia de que ésta se incluya en el proceso de revisión. Resulta de gran ayuda la utilización de textos interculturales donde esta minoría étnica esté presente, ya que puede incrementar su motivación.
- Incidir en que los alumnos comprendan y sean conscientes de que escribir va más allá de los aspectos formales como la ortografía, puntuación... y conseguir que la escritura tenga sentido para ellos (una finalidad y una razón): el para quién y para qué escriben.
- Poner en marcha programas para que los alumnos *conozcan y apliquen estrategias útiles* en la composición escrita en cada uno de los procesos de composición (véase Salvador y García, 2006).

- Trabajar la composición escrita haciendo hincapié en la importancia de los procesos cognitivos: a) partir de una imagen y/o situación dada y discutir en grupo, con la ayuda del profesor, sobre la forma y estructura en que se va a escribir y con el fin de organizar y estructurar las ideas; b) realizar en una situación de pequeño grupo tareas de identificación y corrección de errores en todos los niveles (forma, estructura y contenido); c) aplicar estrategias de autorregulación en diferentes situaciones de escritura (auto-instrucciones, auto-control y auto-refuerzo); d) plantear la utilización de «hojas de pensamiento» (*thinking-sheets*) para promover procesos cognitivos de mayor complejidad.

## Prospectiva de investigación

Según los resultados obtenidos y la extensa revisión bibliográfica realizada sobre este tema, se sugieren algunas prospectivas de investigación:

- Comparar los resultados obtenidos en el proceso de producción de un texto con los que se obtenga en el texto producido en sí (texto que escriben los alumnos), así como conocer cómo la actitud hacia la escritura y la percepción de auto-eficacia pueden influir en el desarrollo de la competencia escritora de estos alumnos. Aspectos que están siendo estudiados en la actualidad como parte de la tesis doctoral.
- Realizar investigaciones similares para conocer si los resultados obtenidos con estos alumnos de etnia gitana también se dan en otros con la misma situación de desventaja socioeducativa.
- Investigar de forma cualitativa cómo se lleva a cabo la enseñanza de la composición escrita en el aula y comprobar en qué procesos cognitivos y meta-cognitivos se hace mayor o menor hincapié.
- Conjugar los resultados obtenidos en los procesos cognitivos individuales con la realización de entrevistas clínicas tanto al profesorado como a los padres, con el objetivo de aportar información valiosa acerca de problemas familiares y escolares que pueden afectar a la habilidad del alumno para escribir.

- Conocer cómo el profesorado influye o puede influir en el aprendizaje de la composición escrita en este tipo de alumnado, desde una triple perspectiva: social, cultural y cognitivo-afectiva.

## Referencias bibliográficas

- BARDÍN, L. (1986). *Análisis de contenido*. Madrid: Akal.
- BEREITER, C. Y SCARDAMALIA, M. (1987). *Psychology of written composition*. New Jersey: Hillsdale Erlbaum.
- BERNINGER, V.W. Y WHITAKER, D. (1993). Theory Based Branching Diagnosis of Writing Disabilities. *School Psychology Review*, 22 (4), 623-42.
- Bueno, J. J. (1993). *El lenguaje de los niños gitanos. Una perspectiva funcional*. Salamanca: Amarú ediciones.
- CAMPS, A. Y RIBAS, T. (2000). *La evaluación del aprendizaje de la composición escrita en situación escolar*. Madrid: CIDE.
- CASTELLÓ, M. (1999). El conocimiento que tienen los alumnos sobre la escritura. En I. I. POZO, Y C. MONEREO (eds.), *El aprendizaje estratégico* (p. 197-217). Madrid: Santillana.
- CUETOS, F., SÁNCHEZ, C. Y RAMOS, J. (1996). Evaluación de los procesos de escritura en niños de primaria. *Bordón*, 48 (4), 445-456.
- DE CASO, A. M. Y GARCÍA SÁNCHEZ, J. N. (2006). ¿De qué carecen los programas de intervención en escritura actuales? La necesidad de instrucción en motivación hacia la escritura. *Estudios de Psicología*, 27(2), 221-242.
- DE BEUGRANDE, P. (1984). *Text production: toward a science of composition*. New Cork: Routledge Flamer.
- EMIG, J. (1971). *The composing processes of twelfth graders*. Urbana, IL: National Council of Teachers of English.
- FLOWER, L. Y HAYES, J. (1981). A cognitive process theory of writing. *College Composition and Communication*, 32, 365- 87.
- FOX, D. J. (1981). *El proceso de investigación en educación*. Pamplona: EUNSA.
- FUNDACIÓN SECRETARIADO GITANO. (2002). *Evaluación de la normalización educativa del alumnado gitano en Educación Primaria*. Madrid: Asociación Secretariado General Gitano.
- GARCÍA, A. (2003). Procesos cognitivos en la composición escrita de niños con deprivación sociocultural. *Enseñanza*, 21, 233-244.

- GARCÍA SÁNCHEZ, J. N. Y FIDALGO, R. (2003). Cambios en la metacognición de los procesos psicológicos de la escritura en estudiantes de 3º E. P. a 3º ESO. *Revista de Psicología General y Aplicada*, 56 (2), 239-251.
- GONZÁLEZ SEIJAS, R. M. (2002). *Poder escribir: programa de entrenamiento en los procesos cognitivos de la escritura. Libro del educador*. Madrid: EOS.
- GOBO, G. (2004). Sampling, representativeness and generalizability. En C. SEALE., G. GOBO., J. F. GUBRIUM. Y D. SILVERMAN. (ed.). *Qualitative Research* (p. 435-456). London: Sage.
- GRAVES, D. H. (1975). An examination of writing processes of seven year old children. *Research in the teaching of English*, 9, 227-241.
- HAYES, J. R. (1996). A new framework for understanding cognition and affect in writing. En C. M. LEVY Y S. R. RANSELL (eds.). *The Science of Writing* (p. 1-27). New Jersey: Erlbaum.
- HERNÁNDEZ, A. (1996). *Estudios experimentales en logopedia*. Granada: Centro socio-cultural gitano.
- KRIPPENDORF, K. (1997). *Metodología de análisis de contenido: Teoría y práctica*. Barcelona: Paidós.
- MACARTHUR, C. A., GRAHAM, S. Y FITZGERALD, J. (2006). *Handbook of Writing Research*. New York: Guilford Press.
- ORIENTE, M.A. (1979). *Cómo aprenden a leer los niños gitanos*. Madrid: S.N.
- PIAGET, J. (1965). *El lenguaje y el pensamiento del niño pequeño*. New Cork: Meridian Book.
- POVEDA, D., CANO, A. Y PALOMARES-VALERA, M. (2005). La escritura vernácula de las niñas y niños gitanos. *Cultura y Educación*, 17 (1), 53-66.
- RICHARDSON, V. (ed.) (2001) (4th edition). *Handbook of research on teaching*. Washington: American Educational Research Association.
- SALVADOR, F. (2000). *Cómo prevenir las dificultades en la expresión escrita*. Archidona: Aljibe.
- (2005). *La expresión escrita de alumnos con necesidades educativas especiales. Procesos cognitivos*. Archidona: Aljibe.
- SALVADOR, F. Y GARCÍA, A. (2005). *Metodología de la Investigación*. En F. SALVADOR (ed.), *La expresión escrita de alumnos con necesidades educativas especiales. Procesos Cognitivos*. (p. 45-62). Archidona: Aljibe.
- (2006). La enseñanza de la composición escrita, basada en estrategias: aspectos conceptuales y de investigación. *Innovación educativa*, 16, 245-256.
- (en prensa). El proceso de revisión en la composición escrita de alumnos de Educación Primaria. *Revista Española de Pedagogía*.

- SAN ROMÁN, T. (1990). *Gitanos de Madrid y Barcelona. Ensayos sobre aculturación y etnicidad*. Barcelona: Universitat Autònoma de Barcelona.
- SMITH, S. (1997). The Genre of the End Comment - Conventions in Teacher Responses to Student Writing (Composition Instruction). *College Composition and Communication*, 48, 249-268.
- STAKE, R. (2005). Qualitative case studies. En N. K. DELZIN Y Y. S. LINCOLN. (ed.), *Handbook of qualitative research* (2<sup>nd</sup> ed.), (p. 443-466). Thousand Oaks, : Sage.
- TEBEROSKY, A. (1992). Aprendiendo a escribir. Barcelona: ICE/Horsori.
- TOLCHINSKY, L. (1993). *Aprendizaje del lenguaje escrito. Procesos evolutivos e implicaciones didácticas*. Barcelona: Anthropos.
- VIGOTSKY, L. S. (1978). The prehistory of written language. En M. COLE y OTROS (eds.), *Mind in Society: The development of higher psychological processes*. Harvard: University Press.
- VILÁ, M. (1998)). Expresión escrita en niños y niñas del ciclo superior de EGB de Gerona. Estudio descriptivo sobre el producto escrito. Un estudio de casos sobre el proceso de composición escrita. Tesis doctoral. Girona, Universidad de Girona.

**Dirección de contacto:** Antonio García Guzmán. Universidad de Granada. Facultad de Ciencias de la Educación. Departamento de Didáctica y Organización Escolar. Campus Universitario de Cartuja, s/n. Granada, España. E-mail: antogagu@ugr.es

# Mujeres agentes de cambio en la dirección de organizaciones universitarias

## Women as change agents in the headship of higher education organizations

Julián López Yáñez  
Marita Sánchez Moreno

*Universidad de Sevilla. Facultad de Ciencias de la Educación. Departamento de Didáctica y Organización. Sevilla, España*

### Resumen

Este artículo aborda la relación entre el liderazgo de las mujeres y su papel en la transformación de las organizaciones universitarias. Para ello tomaremos como base ocho estudios de caso –y de modo particular dos de ellos– analizados en el marco de una investigación más amplia sobre las mujeres que ocupan cargos de gestión en las universidades españolas. Pondremos en relación el contexto organizativo en el que dichas directivas desarrollaron su gestión, los estilos de liderazgo que desplegaron, así como las transformaciones que emprendieron en momentos críticos de las organizaciones que dirigían. Nuestras conclusiones ponen de manifiesto una especial habilidad en dichas directivas para conducir sin traumas estos cambios y habilidad para preservar en buena medida la cohesión social y un ambiente favorable de trabajo. Al parecer estos logros fueron conseguidos mediante estilos de liderazgo bastante flexibles, contruidos sobre la base de un uso alternativo de variadas fuentes de poder, de tal manera que no fue posible identificar algo así como un único y homogéneo «estilo femenino» de gestión. Por el contrario, sí se pudieron identificar ciertos rasgos comunes de liderazgo y algunas trayectorias similares en el acceso a los cargos de gestión. Todo ello nos permitirá discutir el papel innovador que las mujeres desempeñan y podrían desempeñar en la educación superior. En este sentido, una conclusión destaca sobre las demás: aparentemente las estructuras débilmente acopladas, propias de las organizaciones universitarias requieren formas «blandas» de

ejercicio del poder, es decir, no tanto basadas en la estructura formal, como en una trama de redes sociales y comunitarias.

*Palabras clave:* organizaciones educativas, universidad, innovación, liderazgo, poder, acceso a la dirección, estudios sobre la mujer, metodología cualitativa, estudio de casos.

### **Abstract**

This paper tackles the relationship between women's leadership and their role in the transformation of university organizations. It is based on eight case studies analysed in the context of a broader research work on women who occupy high managerial posts in Spanish universities. Two of the cases will be more deeply described and discussed in terms of the relationships in relation to the organizational context in which these female leaders performed their managerial functions, the leadership styles that they deployed, as well as the modifications which they promoted in critical moments of the organizations that they ran. The final conclusions reveal the outstanding ability of these female managers to smoothly drive through such changes, greatly preserving social cohesiveness and an optimal climate in the workplace. This was achieved by adopting a flexible leadership style on the basis of a wide range of power sources. Nothing such a unique and homogeneous 'feminine style of leadership' was found, although some traits and some common trajectories to reach power positions were identified. On the basis of such findings it will be suggested that loosely coupled structures characteristics in higher education organization need 'soft' ways of power management, more oriented towards informal social networks than formal issues. This will allow us to discuss the innovative role that women potentially, and currently, play in higher education organizations.

*Key Words:* school organizations, higher education, innovation, leadership, power, headship access, women's studies, qualitative methodology, case studies.

## **Perspectiva teórica**

Las universidades españolas han experimentado en los últimos años importantes transformaciones estructurales, en paralelo a un aumento significativo de su producción científica y a una mayor proyección internacional de un número significativo de investigadores y equipos de investigación. Estos cambios exigen nuevos papeles en los directivos y gestores y nuevos enfoques de la acción directiva, dado que las características particulares de las organizaciones universitarias reclaman liderazgos sofisticados

en los que el ejercicio de la autoridad o de la gestión burocrática resulta a todas luces insuficiente (Nidiffer, 2001; Fullan, 2002).

Sin duda las mujeres tienen algo muy importante que aportar en la búsqueda de estas nuevas configuraciones del liderazgo, dado que, a pesar de su menor presencia en los cargos directivos, diversas investigaciones muestran un alto nivel de satisfacción hacia el liderazgo ejercido por las mujeres en las organizaciones desde la perspectiva de sus miembros (Erkut, 2001; Acker, 2005). Aparentemente, una de las razones esgrimidas parece ser la ventaja que en general tienen las mujeres en el manejo de una variedad de habilidades comunicativas y sociales, especialmente adecuadas para el funcionamiento más participativo, reticular, descentralizado, emergente y democrático característico de las organizaciones modernas y en especial de las universidades (Bensimon, 1993; Calás y Smircich, 1993; Grogan, 1996; Tierney y Bensimon, 1996; Erkut, 2001; Fletcher, 2001).

Precisamente, la definición del tipo de liderazgo necesario en las organizaciones modernas ha constituido la búsqueda principal de una teoría de la organización que en los últimos años ha abandonado casi definitivamente la idea de las organizaciones como máquinas o burocracias para asumir mayoritariamente la idea de las organizaciones como construcciones sociales (Alvesson y Deetz, 1996; Chia, 2003; Willmott, 2003). Así pues, la incorporación lenta, aunque progresiva, de las mujeres a los puestos de dirección puede constituir un auténtico banco de pruebas para esta teoría de la organización y en especial para la teoría del liderazgo. Además, ello plantea la oportunidad de una alianza táctica entre las tradiciones teóricas feministas y las de estudio de las organizaciones sociales, para el logro de ciertos intereses comunes, como han sugerido diversos autores (Calás y Smircich, 1993; Perreault, 1993; Gherardi, 1995; Yancey Martin y Collison, 2002).

Esta alianza debería permitirnos ir más allá de algunos enfoques que han caracterizado la literatura y la investigación sobre el liderazgo de las mujeres. Una buena parte de esta literatura se ha caracterizado por mostrar el modo en que los valores masculinos han impregnado el ejercicio del liderazgo en las organizaciones (Blackmore, 1999), mientras que otra parte se ha enfrascado en el intento de definir un estilo femenino de liderazgo que se plantearía como alternativo al masculino. Sin embargo, una corriente cada vez más importante sugiere que, pese a que el género constituye una dimensión constitutiva de la acción en las organizaciones, ésta no debería encerrarse en categorizaciones cerradas y homogéneas acerca de lo masculino y lo femenino. Como han sugerido Alvesson y Billing (1992), tanto el liderazgo ejercido por mujeres como el ejercido por hombres es lo suficientemente diverso y contradictorio en la práctica como

para reducirlo a un modelo bipolar. Estos autores opinan que agrupar bajo el rótulo de «liderazgo femenino» un conjunto de conductas que consideramos adecuadas para el desarrollo de las organizaciones, además de no sostenerse sobre una base empírica, podría incluso desincentivar, tanto en hombres como en mujeres, la adopción de dichos rasgos. En consecuencia, ello no ayudaría a romper los estereotipos que han atravesado tradicionalmente este campo (Billing y Alvesson, 2000).

También Court (2005) ha cuestionado la reducción del fenómeno género-liderazgo a los modelos masculino y femenino. Al igual que Alvesson y Billing (1992) y otros autores escépticos ante el constructo «liderazgo femenino», Court sugiere que la dimensión género debe ser puesta en relación con otras dimensiones que influyen sobre el comportamiento de los líderes, como por ejemplo la clase social o las creencias y la cultura.

En definitiva, los estudios acerca del género en las organizaciones comienzan a superar los enfoques iniciales de las pasadas dos décadas que se centraban en la diferenciación de estilos y características atribuidas a hombres y mujeres. Los nuevos estudios utilizan las aportaciones de una variedad de disciplinas, y toman en consideración tanto las variables situadas en un nivel macro, es decir las condiciones sociales, culturales y políticas que influyen en el liderazgo (Blackmore, 2005), como las variables situadas en el nivel medio o institucional, por ejemplo la historia institucional, la estructura o la cultura organizativa (Mills, 2002). Estos estudios tienden a adoptar un enfoque más procesual y metodologías basadas en el estudio de casos para explorar las historias de líderes concretos y comprender el modo en que todas estas influencias dan forma a su liderazgo (Bensimon, 1993; Collard y Reynolds, 2005).

A pesar de todo, hay un punto de convergencia entre las investigaciones sobre el papel de las mujeres en las organizaciones y la literatura sobre el liderazgo, el cual tiene que ver con el uso de las habilidades sociales y un enfoque centrado en las personas. En este sentido, el giro hacia los aspectos sociales de la teoría de la organización converge con una buena cantidad de estudios empíricos que señalan que aquellos aspectos, sin ser exclusivos de las mujeres, están con mucha frecuencia presentes en sus estilos de liderazgo (Erkut, 2001; Fletcher, 2001).

La búsqueda de liderazgos distribuidos, dialógicos y centrados en las personas es particularmente necesaria en las organizaciones universitarias, donde las diversas formas de gerencialismo se han mostrado especialmente ineficaces en la creación de culturas innovadoras y creadoras de desarrollo sostenible de la docencia y la investigación (Blackmore, 2005). En ellas adquiere aún más sentido la pregunta acerca de cuál es y cuál podría ser la aportación de las gestoras universitarias y directoras de departamentos

universitarios en la construcción de modos colegiales de trabajo mediante liderazgos sostenibles, *visibles* y reconocidos. En la misma línea, cabe también preguntarse si las mujeres que ocupan puestos de dirección y gestión en la universidad pueden representar un impulso a la transformación de éstas. Esta pregunta nos parece especialmente importante en unos momentos en los que –como sostiene Acker (2005)– la incorporación de las mujeres a las tareas de gestión coincide con un aumento significativo de la complejidad que afrontan dichas tareas. En este sentido, nuestro estudio trató de poner en relación algunas características del liderazgo que las mujeres participantes desplegaron mientras se encontraban inmersas en procesos de cambio en sus unidades académicas. Por ello, los estudios de caso que realizamos trataron de describir en profundidad las culturas organizativas y las redes de poder establecidas en dichas unidades. Nos interesaba particularmente indagar acerca de la relación existente entre las estrategias que pusieron en juego estas directivas y la cultura instalada en las instituciones ya que, como señalan Kezar y Eckel (2002), cuando las estrategias de cambio violan las normas culturales es difícil que se institucionalicen los cambios acometidos.

## Metodología

El papel que las directivas desempeñan en la transformación de las organizaciones universitarias en este momento de profundas reformas en la educación superior fue uno de los objetivos de una investigación más amplia que realizamos entre 2002 y 2005 acerca de las mujeres que desempeñan tareas directivas y de gestión en universidades españolas (como rectoras, vicerrectoras, decanas, vicedecanas, directoras y secretarías de departamento). Esta investigación se propuso además conocer los estilos de liderazgo y la forma en que estas directivas ejercen el poder en sus organizaciones.

Se desarrolló en dos fases. La primera tuvo un carácter descriptivo y se basó en los datos obtenidos a partir de dos cuestionarios diseñados *ad hoc*. El primero de ellos fue respondido y devuelto por 136 directivas que representaban el 31,26% de las 435 que componían la muestra de mujeres que en aquel momento desempeñaban cargos de gestión, a todas las cuales les fue enviado. Indagaba acerca del perfil profesional de las participantes en el estudio, los papeles desempeñados, sus estilos de liderazgo, los problemas y las necesidades formativas percibidos por ellas, así como su contribución a la construcción de la cultura de sus unidades académicas. El segundo cuestionario estuvo

dirigido a las personas que trabajaban en las unidades dirigidas por las participantes en nuestro estudio. Su propósito fue contrastar la visión que de sí mismas ofrecieron las directivas en el primer cuestionario, en particular respecto a la forma de ejercer el poder, así como analizar el impacto que dicho ejercicio estaba teniendo sobre el funcionamiento institucional de dichas unidades. Para ello nos concentramos en los departamentos, que eran las unidades que presentaban mayor interés para el mencionado propósito. Hicimos una selección de éstos basada en los siguientes criterios: distribución equilibrada por comunidades autónomas, por tipo de Universidad, de acuerdo con su gestión pública o privada, por área de conocimiento del departamento, y en función del estilo de liderazgo de la directiva correspondiente. Este cuestionario fue respondido por 91 profesores, pertenecientes a 17 de los 19 departamentos consultados.

Las conclusiones más destacables de esta primera fase fueron: (a) Ya en el establecimiento de la población del estudio se evidenció el desequilibrio en el acceso a los cargos de gestión por parte de profesoras y profesores a favor de estos últimos. (b) Se identificó la tendencia predominante hacia un estilo de liderazgo flexible y adaptativo que utiliza una amplia gama de fuentes de poder en función de las exigencias de cada situación. (c) Se identificó el predominio en las directivas de una *visión policrónica* de la gestión y la resolución de problemas. (d) El interés por el cuidado de las relaciones sociales y la atención a las necesidades de los miembros de la organización. (e) El deterioro de la dedicación a las tareas de investigación como principal consecuencia del ejercicio de cargos de gestión.

En la segunda fase, de carácter cualitativo, se seleccionaron ocho casos para su estudio en profundidad, sobre la base de diferentes criterios. El primero de ellos fue la orientación predominante del estilo de liderazgo, obtenido a partir de las respuestas al primer cuestionario. Para ello se cruzaron los componentes (a) *autoridad y planificación* y (c) *orientación hacia las relaciones sociales*, asignándosele a ambos valor positivo o negativo. Previamente se había descartado utilizar el componente (b) *orientación ética y responsabilidad*, ya que una amplia mayoría de las encuestadas aparecía con valor positivo en este componente. De dicho cruce de componentes resultaron cuatro modelos teóricos, definidos en función de si las encuestadas consideraban que contaban con: (1) *condiciones de ambas dimensiones*, (2) *condiciones de la dimensión A pero no de la C*, (3) *condiciones de la dimensión C pero no de la A*, (4) *no contaban con condiciones de ninguna de las dimensiones definidas*. El uso de la función de *formación de conglomerados de K-medias* del paquete estadístico SPSS 11.0, nos permitió identificar 18 casos que podían asignarse a alguno de los cuatro modelos teóricos, tal como puede verse en la Tabla I.

TABLA I. Fase cualitativa. Selección inicial de los casos

Habilidades sociales e interés por las relaciones interpersonales (C)	C+	<b>50 Científica</b> <b>404 Jurídicas y sociales</b> <b>335 Científica</b>	<b>167 Humanidades</b> <b>325 Humanidades</b> <b>92 Jurídicas y sociales</b> <b>24 Humanidades</b> <b>156 Jurídicas y sociales</b> <b>253 Científica</b>
	C-	<b>259 Técnica</b> <b>55 Técnica</b> <b>179 Científica</b> <b>141 Jurídicas y sociales</b> <b>8 Jurídicas y sociales</b> <b>385 Jurídicas y sociales</b>	<b>176 Científica</b> <b>84 Científica</b> <b>130 Jurídicas y sociales</b>
		A -	A +
		Autoridad y habilidades de gestión y planificación (A)	

A partir de aquí se realizó una segunda selección, esta vez *manual*, buscando un cierto equilibrio en los criterios siguientes: comunidad autónoma, área de conocimiento, gestión pública o privada de la universidad y cargo de la directiva. También se consideró que el volumen de respuestas obtenidas al segundo cuestionario –el que cumplimentó el resto de profesores del departamento en cuestión– fuera suficientemente amplio. Finalmente, la selección quedó constituida por dos casos pertenecientes a cada uno de las cuatro modelos teóricos, como puede verse en la Tabla II (los nombres de las directivas son ficticios).

TABLA II. Fase cualitativa. Selección definitiva de los casos

Orientación liderazgo	Caso	Área	Tipo de Universidad	Cargo
1 A + C +	Diana	C, Jurídicas y Sociales	Pública	Decana
	Berta	Humanidades	Privada	Directora Dpto.
2 A - C +	Enma	C, Jurídicas y Sociales	Pública	Vicedecana
	Irene	Científica	Pública	Ex Directora Dpto.
3 A - C -	Pilar	Científica	Pública	Directora Dpto.
	Ofelia	C, Jurídicas y Sociales	Pública	Directora Dpto.
4 A + C -	Leonor	Científica	Pública	Ex Directora Dpto.
	Ana	C, Jurídicas y Sociales	Pública	Vicedecana

La información para el análisis de los casos se obtuvo a través de entrevistas, observación de campo y de reuniones y estudio de documentos. Respecto a las entrevistas, éstas se realizaron tanto a la directiva protagonista como a algunos de los seguidores, incluyendo al menos a una persona consolidada en la institución, una persona joven en la institución, un miembro del personal de la administración y los servicios y, en algunos casos, a estudiantes. Las entrevistas se realizaron siguiendo guiones semi-estructurados que, en el caso de las directivas, se referían a su trayectoria personal y profesional, cuestiones relacionadas con el desempeño del cargo (experiencias y percepciones, prioridades y preocupaciones, estilo u orientación, problemas encontrados, condicionantes en el acceso y necesidades formativas), valoración de su gestión, así como cuestiones sobre aspectos institucionales: proyectos realizados, acontecimientos significativos, conflictos, relaciones con otras entidades, historia institucional, cambios experimentados, etc. Por otra parte, el guión de las entrevistas dirigidas a los seguidores incluía preguntas acerca del papel que desempeñaban en la organización y de su trayectoria en ella, acerca del desempeño en la gestión de la directiva protagonista del caso, sus características como gestora y estilo de liderazgo, y acerca de las condiciones organizativas en las que desarrolla su trabajo (satisfacción personal, problemas relevantes, oportunidades de mejora, proyectos e iniciativas, ambiente de trabajo).

Además se analizaron documentos (actas de consejos de departamentos, comunicaciones formales, agendas, proyectos en curso, etc.) y se observaron actos de gestión, como consecuencia de un acompañamiento sistemático de la directiva protagonista durante un período de tiempo negociado con ella, en las actividades programadas en su agenda como, por ejemplo, juntas de facultad, consejos de departamentos, reuniones del equipo directivo, etc. También los encuentros informales mantenidos con las entrevistadas, pese a no ser registrados, proporcionaron una rica información sobre su estilo de liderazgo. Las visitas tuvieron una duración que osciló entre un día y tres, de acuerdo con la disponibilidad de los miembros de cada organización.

Las entrevistas fueron grabadas en audio y posteriormente transcritas para su análisis. En dicho análisis se utilizó el sistema de categorías que se muestra en la Tabla III, tanto para las entrevistas a las directivas como a los *seguidores*. Se empleó una codificación múltiple: descriptiva e interpretativa y se hizo uso del programa de ordenador AQUAD 5 para el procesamiento y análisis del conjunto de la información.

TABLA III. Sistema de categorías para el análisis de las entrevistas

<p><b>0. TRAYECTORIA PERSONAL Y PROFESIONAL</b>  Información biográfica de la persona entrevistada, a fin de poner en relación sus circunstancias personales con el desarrollo de su vida profesional.</p>
<p><b>1. ACCESO AL CARGO</b>  Esta categoría engloba la información mediante la que se describe el momento y las razones del acercamiento a las labores de gestión, por parte de la protagonista del caso.  1.1 CIRCUNSTANCIAS DE ACCESO AL CARGO  1.2 MOTIVACIONES EN EL ACCESO AL CARGO.</p>
<p><b>2. PERCEPCIONES SOBRE EL GÉNERO</b>  La categoría reúne las opiniones relativas a la incidencia de la variable género en la vida profesional de la protagonista.  2.1 EXPERIENCIAS RELATIVAS A LA INFLUENCIA DEL GÉNERO EN EL DESARROLLO PROFESIONAL.  2.2 PERCEPCIONES SOBRE LA INFLUENCIA DEL GÉNERO EN EL DESARROLLO PROFESIONAL.</p>
<p><b>3. DESEMPEÑO DEL CARGO/ TAREA</b>  Actividades de gestión que desempeña la protagonista, así como del modo en el que las lleva a cabo y el impacto que tienen en ella.  3.1. NECESIDADES FORMATIVAS.  3.2. PERCEPCIONES ACERCA DE LOS ASPECTOS DE LA TAREA QUE PRIORIZA.  3.3. PERCEPCIONES SOBRE LOS ASPECTOS DE LA TAREA DESCUIDADOS.  3.4. PREOCUPACIONES SOBRE EL CUMPLIMIENTO DE LA TAREA («EL ESPACIO EUROPEO»);  3.6. SENTIMIENTOS EN RELACIÓN CON LAS ACTIVIDADES DE GESTIÓN.  3.7. ACTITUDES Y ESTILO DE LIDERAZGO.</p>
<p><b>4. PROBLEMAS DETECTADOS EN RELACIÓN CON LA TAREA.</b>  Dificultades de diverso orden que la gestora ha encontrado –según su opinión o la de quienes trabajan con ella– para desempeñar tal función.  4.1 EN RELACIÓN CON LOS COMPAÑEROS Y LOS GRUPOS.  4.2 RELATIVOS A LA INSTITUCIÓN / DEPARTAMENTO.  4.3 DE ORDEN INDIVIDUAL: RELACIONADOS CON LA ORGANIZACIÓN DE LA TAREA, CON EL EMPLEO DEL TIEMPO Y/O PERSONALES, Y CON EL EJERCICIO DEL CARGO.</p>
<p><b>5. OPINIONES GENERALES</b>  Expresa las opiniones, tanto de la gestora como de las personas entrevistadas como contraste, acerca del liderazgo como función y de la influencia de diversas variables en su ejercicio.  5.1 ACERCA DEL GÉNERO.  5.2 SOBRE EL LIDERAZGO Y LA FUNCIÓN DIRECTIVA: CARACTERÍSTICAS DE UN BUEN GESTOR; CONOCIMIENTOS VALORADOS COMO NECESARIOS.</p>
<p><b>6. HISTORIA Y CONDICIONES INSTITUCIONALES</b>  Se refiere a las características institucionales que, de acuerdo con la posición asumida en la investigación, pueden incidir en el desarrollo organizativo y en el ejercicio del liderazgo y las funciones de gestión en la Universidad particular. Contempla: <i>Proyectos, Conflictos organizativos, Relaciones institucionales con diferentes instancias, etc.</i></p>

Los informes de cada caso se realizaron por parejas de investigadores. A continuación, dicho informe fue revisado por el resto del equipo de investigación, como medio de control de la implicación de los investigadores en relación con su objeto de estudio. Una vez se hubo elaborado el informe de cada caso, se realizó un análisis transversal para establecer regularidades entre ellos y conclusiones a la luz de la literatura específica. A continuación se expondrán algunos aspectos de dos de los ocho casos estudiados en esta segunda fase de la investigación. Pensamos que ambos casos ilustran de forma muy significativa el modo en cómo las gestoras universitarias estaban liderando procesos de cambio en sus respectivas instituciones.

## Diana y la Facultad de Ciencias Sociales

Diana es Decana en una relativamente pequeña y joven -27 años- universidad pública. Dicha Universidad ha experimentado notables transformaciones en los últimos años. En particular, ha crecido notablemente, se ha trasladado a un nuevo y moderno campus y ha descentralizado sus estructuras y servicios administrativos, lo cual ha dado más poder a los decanos, en detrimento de los directores de los departamentos.

Diana es profesora de esta universidad desde 1990. Durante los primeros años como profesora debió compaginar el desarrollo de su carrera académica con la puesta en marcha de las asignaturas de su perfil docente, las cuales se vieron modificadas muy pronto por la implantación de un nuevo plan de estudios, lo que supuso una importante tarea añadida. Sin embargo, a pesar de ello, Diana encontró de forma muy prematura tiempo y ganas -desde 1993- para desempeñar actividades de gestión. Se trataba al principio de cargos honoríficos, hasta acceder en 1999 a la Secretaría de su Departamento. No fue ésta, sin embargo, una decisión exclusivamente personal, sino que formó parte de un plan concebido por un grupo de mujeres de su Departamento -que siempre había estado dirigido por hombres- para optar progresivamente a diversos cargos, tanto en el Departamento como en la Facultad, que habían quedado o iban a quedar vacantes. Muy pronto este «lobby» o grupo de poder daría sus frutos y Diana llega a ser Vicedecana en un equipo de gobierno de la Facultad en la que la decana es Pina, otro destacado miembro del mismo. Más adelante, cuando Pina recibe el ofrecimiento de ocupar la Vicerrectoría de Relaciones Internacionales, la cual acepta y ocupa hasta el presente, Diana es elegida Decana con un equipo que trata de dar continuidad a los proyectos iniciados en la etapa anterior y en el que también están presentes algunos hombres.

### Una situación difícil

Como para muchas otras jóvenes académicas, los cargos de gestión supusieron para Diana la oportunidad de integrarse rápidamente en un contexto ocupado mayoritariamente por hombres, aunque ello le acarrearía importantes esfuerzos de conciliación de sus variadas ocupaciones con la vida familiar. Pero Diana encontraría dificultades añadidas. Ella asumió la dirección de la Facultad en un momento delicado, dado que acababa de producirse la división formal de su Departamento en dos nuevos departamentos, algo que se había estado preparando durante largo tiempo y que estuvo precedido por una etapa de conflictos en el marco de una fuerte micro-política.

Diana había sido elegida como Decana siendo la candidata del grupo (área de conocimiento) mayoritario en el departamento anterior a la división. Su propuesta como candidata se justificó por el hecho de que había sido una de las personas menos significadas en las luchas internas y, por tanto, podía granjearse un cierto reconocimiento también por parte de «la oposición». El líder de esta oposición reconoció que (con Diana) «las cosas se han tranquilizado», que «se ha recuperado la posibilidad de hablar y de hacer cosas juntos» y que «la situación va camino de cambiar». También que Diana «quiere hacerlo bien» y que una de sus mayores preocupaciones es «ser aceptada por todos».

### **Poder en la sombra**

Uno de los problemas principales señalados por Diana en su acción de gobierno resultó de gran interés para nuestro estudio. Su equipo decanal estaba compuesto al 100% por profesores –mujeres y hombres– titulares o asociados, con la total ausencia de los catedráticos (todos ellos varones). Según Diana, los hombres «siguen dominando por detrás, porque son catedráticos, porque son lo que fueron, porque tienen un prestigio, pero sobre todo porque son catedráticos». Los investigadores pudieron comprobar esta pauta implícita de dominio desplegada por «los catedráticos» en una reunión de una Junta de Facultad a la que se nos permitió asistir como observadores. En dicha reunión, este grupo se valió de una curiosa estrategia para el mantenimiento de una posición hegemónica: el humor. Los catedráticos, en un ambiente distendido, acapararon la mayor parte del tiempo disponible de reunión mediante intervenciones verbales en las que de manera muy frecuente intercambiaban bromas entre sí. De hecho se convirtieron en los protagonistas de la reunión, debiendo realizar la Decana esfuerzos constantes para cortar sus digresiones y volver a los temas señalados en el orden del día. No es extraño que Diana sitúe en los privilegios adquiridos por los profesores catedráticos, especialmente los relacionados con la elección de horarios/asignaturas, algunos de los problemas más importantes que encuentra en su gestión.

### **Estilo de liderazgo de Diana**

Diana basó su gestión desde el primer momento en el trabajo en equipo, por eso el proceso de formación de su equipo de gobierno fue muy importante para ella. El criterio que utilizó fue muy diferente al de la decana anterior, que había elegido a sus colaboradores

tratando de mantener un siempre difícil equilibrio entre los departamentos y áreas de conocimiento presentes en la facultad. Para Diana, por el contrario, lo prioritario era la complicidad y la armonía entre los miembros del equipo. La importancia concedida a estos aspectos hace que se disguste seriamente ante el abandono de uno de sus vicedecanos para ocupar otro cargo de gestión. También lo demuestran las condiciones con las que ha tratado de crear un clima de trabajo satisfactorio. Por ejemplo, el equipo tiene reuniones sistemáticas todos los miércoles, las cuales comienzan con café y galletas o pasteles que sus miembros aportan. Con cierta frecuencia almuerzan juntos y en alguna ocasión se han «encerrado» a trabajar durante una jornada completa en una apartada y tranquila casa rural que han alquilado para la ocasión.

Pero además, Diana mantiene una amplia red de colaboradores informales cuya importancia en dicha red no está directamente determinada por la posición que esas personas ocupan en el organigrama. Por ejemplo, Marta, una de las responsables administrativas de la Facultad, es una pieza importante de dicha red. Se trata de la responsable del área de alumnos, y no es casualidad que la atención a los alumnos y la escucha frecuente a los representantes de éstos sea uno de los ejes de su gestión. Además Marta conoce a fondo esta Universidad casi desde sus inicios, conocimiento que también adquiere una especial importancia para el modo de gestión reticular e informal que adopta Diana. Pero la red se extiende también a otros profesores de su Departamento o de la Facultad que no forman parte de su equipo, a estudiantes, a otros miembros de la administración y, por supuesto, a miembros del equipo rectoral, como Pina, con la que mantiene inmejorables relaciones.

Además de la importancia concedida al trabajo en equipo y a la construcción de una amplia y difusa red de colaboradores, el estilo de gestión de Diana se caracteriza también por su preocupación por integrar hasta el límite de lo posible las diferentes sensibilidades que existen dentro de la facultad. Esto significa desplegar un programa con objetivos claros y un estilo de dirección firme en determinados aspectos, pero a la vez flexible en otros, precisamente para evitar volver a situaciones conflictivas. Sin embargo, Diana asume que hay momentos en los que la búsqueda de consenso se topa con barreras infranqueables, requiriéndose entonces decisiones que, según ella, los directivos deben adoptar con firmeza y convicción, puesto que «la máxima satisfacción es ver que algo sale adelante». Encontramos que es precisamente esta combinación de firmeza y cierta directividad, por un lado, y de amplio despliegue de relaciones sociales y diplomacia, por el otro, la que caracteriza esencialmente el estilo de gestión de Diana. El componente directivo o de autoridad en su estilo de gestión fue puesto de manifiesto por diferentes entrevistados: «ella es firme en sus convicciones»;

«tiene una forma de trabajar más directa»; «a mí me dicen que soy un poco mandona» (la propia Diana). El resultado es un estilo «activista» y «resolutivo» –«eficaz» dirá uno de los entrevistados– que no espera mucho tiempo con los problemas sobre la mesa, sin resolver y no teme adoptar decisiones arriesgadas siempre que cuente con el apoyo de su propio equipo.

## Leonor y el Departamento de Ecología

Leonor lleva casi veinte años como miembro de un departamento constituido por alrededor de veinte profesores, en una Universidad pública de una región periférica. Ha ocupado la dirección del departamento durante cuatro años. Hace apenas uno que abandonó el cargo, al finalizar el segundo de los períodos de dos años para los cuales resultó elegida por la mayoría de sus compañeros. Antes de ser Directora, sin embargo, ya había desempeñado el cargo de Secretaria del Departamento.

Leonor fue precedida en la dirección por Pura (que ahora acaba de re-assumir el cargo) y ésta, a su vez, por Leandro, el fundador del Departamento y director del mismo durante nada menos que 17 años. Pura y Leonor han ejercido el liderazgo de manera muy diferente a Leandro –a pesar de la identificación de ambas con su figura– el cual adoptó un estilo fuertemente carismático y, en buena medida, autocrático, lo que resulta muy interesante para los propósitos de nuestra investigación.

## La marca de la historia institucional

Cuando se constituyó el Departamento de Ecología, en 1978, Leandro era el único catedrático del mismo, así que fue nombrado de oficio por las autoridades. Luego afrontó varias elecciones en las que obtuvo la mayoría hasta abandonar el cargo en 1995 por enfermedad. Los testimonios recogidos rodean a Leandro de un halo heroico: «dio su vida al departamento», «era una persona muy enérgica, muy transgresora, que le infundió vida al departamento», «era capaz de arrastrar», «desde que se fue no se ha vuelto a hacer nada en común», «ninguno tenemos la capacidad de liderazgo que él tenía...» Leandro era capaz de atraer proyectos relevantes al departamento y le dio un impulso decidido a la investigación. La gran mayoría de los profesores se implicaron intensamente

en estos proyectos fundacionales, dándose en el grupo una motivación y un sentimiento de pertenencia muy altos.

En los primeros momentos de su mandato, Leandro apoyó la llegada al departamento de Leonor, Pura y otros profesores jóvenes, con el apoyo de los cuales construye su propio «imperio». Este imperio sólo se desbarataría años más tarde, tras la enfermedad irreversible de Leandro y su baja del departamento. A partir de ese momento, la cohesión que su figura logró se debilita y sus colaboradores compiten con los opositores por la hegemonía. Una de las pugnas más importantes tiene lugar cuando ambos grupos apoyan a sus líderes respectivos para conseguir una plaza de catedrático que es convocada. A pesar de todo, el grupo de los seguidores de Leandro sigue siendo el mayoritario y por eso Leonor y Pura se alternan en la ocupación del cargo durante los años posteriores. En todas las elecciones obtuvieron la mayoría absoluta de los votos del Consejo de Departamento.

Pese a lo identificadas que siempre se mostraron Leonor y Pura con la figura académica y con el perfil personal de Leandro, ellas han actuado en todo momento de manera muy diferente como directoras. Ambas han estrechado las redes de relación social entre sus colaboradores y, aunque las relaciones con los opositores no son cordiales, al menos han conseguido que las reuniones del Consejo del Departamento se desarrollen sin dramas y con diálogo, aunque se discrepe. Precisamente en esta etapa la discrepancia está visible, a diferencia de la etapa de Leandro, donde no lo estaba, aunque existía. Sin embargo, algunos se quejan de que la otrora férrea cohesión impuesta o lograda por el fundador se ha convertido ahora en una amalgama de pequeños *reinos de taifas*, o células constituidas por los diferentes grupos de investigación (cuatro grupos), cada uno de los cuales está compuesto por un reducido número de profesores (entre tres y cinco). Estos grupos cuentan con bastante autonomía, al estar la mayor parte de sus proyectos de investigación financiados por fondos externos que son gestionados por el propio grupo.

Entre tres de estos grupos hay una cierta afinidad en relación con sus intereses académicos, sintiéndose de algún modo sucesores del fundador. La postura del cuarto grupo es, por su parte, abiertamente disidente respecto a la cultura dominante. Este grupo no parece encontrar «su hueco» en el departamento: «hay otras líneas que ellos no quieren que prosperen...». Aunque sus posiciones críticas hacia la dirección del departamento son matizadas («admitiendo que no sean malas [las directoras]»), se muestran disconformes con el monopolio en la dirección de la cultura dominante («encuentro que es muy malo que se estén repitiendo los mandatos»... «puede haber otras maneras»... «en la dirección debe haber mucha apertura, que sea muy cambiante, que todos nos involucremos»).

Tanto Pura como Leonor tienen a gala que no imponen nunca una decisión, que no actúan «piramidalmente». La mayor parte de sus compañeros lo reconoce, sin embargo, algunos se quejan de una gestión «excesivamente formal». Otros alegan que adoptan un perfil meramente administrativo y que el departamento no participa en las decisiones importantes en el ámbito de la facultad o de la universidad. Leonor reconoce que dedica la mayor parte del tiempo al papeleo, pero alega que es inevitable.

Lo que más resaltan sus partidarios es que Leonor y Pura han evitado, tras la marcha de Leandro, que los conflictos fueran a más y que «la sangre llegara al río». Parece ser cierto porque la división no ha trascendido a los recién llegados, los cuales se llevan muy bien entre ellos, con independencia del grupo donde están adscritos. En todo caso, tras la época turbulenta de la refundación, el departamento ha ido recuperando la calma poco a poco y, salvo algunas dificultades aisladas, la gestión del mismo se desarrolla actualmente sin sobresaltos ni graves problemas.

### **Estilo de liderazgo de Leonor... y de Pura**

De algún modo, los liderazgos sucesivos de Pura y Leonor en la dirección del departamento han debido tener un papel fundamental en la construcción de una trama social de sostén que ha impedido que los conflictos generados en el proceso fundacional y durante la transición hayan creado un ambiente insostenible. Pura, la directora actual, habla del tipo de liderazgo que se hizo necesario tras la marcha del fundador, en un departamento muy diverso, que ha exigido a las dos directoras el estar muy atentas a las diferentes sensibilidades de los grupos que lo componen.

Podemos ver la combinación de estilos de ambas directivas como una auténtica estructura de poder institucionalizada que se ve reforzada tanto por el apoyo que les presta la mayoría de los miembros del departamento, como por la débil y «deportiva» presión que ejerce la minoría en la oposición. Ellas apenas ejercen su influencia desde la autoridad, desde la jerarquía. Por el contrario operan combinando ciertos controles burocráticos con un amplio conocimiento e intervención sobre las redes sociales y los grupos informales. Y ello a pesar de que esta especie de renuncia a la autoridad formal representa a veces una dificultad, puesto que las formas «blandas» de poder son interpretadas como debilidad por algunos miembros, en opinión de Pura.

Por tanto, estamos más bien ante una estructura de poder de base meritocrática/burocrática con un importante componente micropolítico, en la medida en que los grupos de investigación funcionan además como coaliciones que negocian de

manera informal la agenda e incluso las decisiones a adoptar en las reuniones formales. Es sobre estas bases de poder que Leonor y Pura han evitado parte de las tendencias disgregadoras características de las estructuras de poder fuertemente politizadas. Parecen haber ayudado a esta contención de la escalada de los conflictos algunos de los valores asumidos por estas mujeres en el ejercicio de la dirección, especialmente el compromiso («queremos al departamento») y la honradez, entendida como una modalidad de compromiso ético. Y también han jugado un papel importante ciertas dosis de diplomacia y el uso intensivo de las relaciones sociales y del diálogo, en lugar de aplicar el rodillo de la mayoría en cualquier circunstancia.

## Conclusiones y discusión

Como ya dijimos, los dos casos expuestos han sido seleccionados para los propósitos de este artículo porque nos parece que ilustran adecuadamente el papel que las directivas universitarias están jugando en el acompañamiento y, en ocasiones, en el impulso de procesos de transformación de sus organizaciones. La fase cualitativa de nuestra investigación puso de manifiesto una especial habilidad en dichas directivas para conducir sin traumas estos cambios y preservando en buena medida la cohesión social y un ambiente favorable de trabajo. Al parecer estos logros fueron conseguidos mediante estilos de liderazgo bastante flexibles, contruidos sobre la base de un uso alternativo de variadas fuentes de poder. En las líneas siguientes trataremos de fundamentar estas conclusiones generales no sólo sobre las evidencias ofrecidas por estos dos casos (Diana y Leonor), sino aludiendo también a los otros seis.

### Circunstancias del acceso al poder

En términos generales podemos afirmar que las directivas de nuestra investigación accedieron a su cargo cuando la organización afrontaba cambios importantes y/o circunstancias difíciles. Leonor asumió la dirección de un departamento fuertemente marcado por la personalidad del director anterior, pero amenazado a su vez por las tensiones disgregadoras que explotaron tras la salida de éste. Diana asumió su cargo de Decana en el marco de importantes transformaciones de su Universidad y en medio de

fuertes tensiones entre el profesorado, las cuales provocaron la división en dos del único Departamento que tenía originariamente la Facultad. En el resto de los casos estudiados, las directivas protagonistas vivieron situaciones similares. Enma vivió la escisión de su Facultad al inicio de su carrera, escisión de la cual salió la nueva Facultad en cuya dirección participa como vicedecana. A Irene le tocó reunir en su departamento, como directora, a dos áreas de conocimiento que estaban vinculadas a departamentos diferentes. Ofelia ha dirigido un departamento del que ha sido prácticamente la fundadora en el marco de una facultad muy joven (6-7 años) y de un centro universitario que ha adquirido ese carácter hace poco.

En ningún caso las directivas analizadas accedieron a cómodos cargos de gestión. Por el contrario, todos ellos revestían una gran complejidad y requirieron una fuerte implicación, al mismo tiempo que pusieron a prueba sus habilidades como gestoras, e incluso como personas. En este sentido nuestros hallazgos coinciden con los de Acker (2005) quien analizó casos de mujeres líderes de facultades universitarias procedentes de tres países diferentes y encontró que dichas mujeres afrontaron dos importantes desafíos: ser parte de la primera generación de mujeres -pocas- implicadas en la gestión universitaria y trabajar en un contexto de tendencias globales que han hecho el trabajo de académicos y gestores más difícil que en épocas pasadas.

Además, las mencionadas dificultades institucionales se añadieron a las del desarrollo profesional y académico. Algunas directivas compaginaron sus cargos con las fuertes exigencias que la Universidad impone en los momentos iniciales de dichas carreras, y todas con la dedicación a las tareas familiares, atravesando por momentos cruciales de indefinición y conflicto de rol (Lea y Leibowitz, 1992; Erkut, 2001; Wolf-Wendel y Ward, 2003). En dos casos -Pilar y Ana- ocurría que las carreras en las que se habían formado no eran las mismas que aquellas en las que iniciaban su andadura como profesoras universitarias. Ello implicó que su iniciación académica fue aún más costosa de lo normal y plagada de las incertidumbres propias del hecho de sentirse extrañas en aquel lugar y de los recelos de sus compañeros. En otro caso, el de Berta, la procedencia era de otro país y las incertidumbres provenían de la necesidad de adaptación a una cultura diferente.

Precisamente porque las circunstancias de acceso al cargo revistieron todas estas dificultades, sorprende en cierto modo constatar que las protagonistas de nuestros casos comenzaron a ejercer cargos de gestión muy pronto desde su incorporación a la universidad. Nada menos que seis de ellas comenzaron a asumir cargos académicos precisamente cuando debían compaginarlos con sus esfuerzos para adquirir la situación de funcionarias. Tan sólo Irene y Pilar eran ya catedráticas y estaban, por lo tanto, en una posición consolidada al acceder a la dirección de sus respectivos departamentos.

En este sentido, desde nuestro punto de vista, la necesidad de aceptación y de integración en el grupo aparece como una de las motivaciones importantes en el acceso a cargos de gestión. Quizás esta necesidad se vea acuciada en el caso de las mujeres debido a su percepción de las organizaciones formales, especialmente las burocráticas, como lugares hostiles o alejados de su forma de ver el mundo y de abordar los problemas (Hargreaves, 1996).

Por otro lado, este primer acceso a un cargo directivo viene en un buen número de ocasiones caracterizado bajo la fórmula de sucesión o acompañamiento de un hombre que había dejado una huella o fuerte impronta en la vida del departamento o de la facultad en cuestión. Berta, por ejemplo, comienza su andadura en la gestión como Jefa de Estudios junto a un director del Departamento muy ocupado, que la necesita para descargar sobre ella ciertas tareas. Enma se postula a vicedecana por su buena relación con el decano, en un momento en el que cree que debe implicarse más en su centro. Leonor apoya primero la gestión del «fundador» del departamento y luego, cuando éste se retira, ocupa su cargo; en parte continuando su legado y en parte introduciendo novedades importantes tanto en la forma como en el contenido de su gestión. En contraste, encontramos un solo caso, el de Diana, en el que la directiva «funda» –digámoslo así– un proyecto propio, como parte de un grupo en el que participan otras mujeres y que imprime un nuevo orden en la organización.

Por lo tanto, aquí encontramos lo que parece ser un modelo en el proceso histórico de las mujeres en el acceso a puestos de gestión en una organización como la universitaria en la que el estatus de catedrático –el rango más alto– está desigualmente repartido entre hombres y mujeres, aunque la tendencia se encamina poco a poco a neutralizar dicha desigualdad. El proceso se inicia cuando algún varón reclama a alguna mujer que la acompañe como escudero en el proceso de gestión. Esta mujer se encuentra en los inicios de su carrera universitaria. La mujer acepta, animada por las posibilidades de una mayor integración en su unidad, y paulatinamente encuentra que sus características personales –y también probablemente su condición femenina– se acoplan especialmente bien a las tareas y funciones que se le reclaman. Con el tiempo su gestión se independiza de la de su mentor y adquiere el protagonismo como directiva.

## **Las mujeres y la organización universitaria ¿nacidos el uno para el otro?**

Ahora bien, la fase cualitativa de nuestra investigación mostró que, una vez llegadas al poder, las mujeres parecen lograr cambios significativos, percibidos con satisfacción

por una mayoría de los miembros de la organización. En concreto, las directivas analizadas se mostraron hábiles y eficaces por lo general en la gestión de la conflictividad latente o manifiesta de sus respectivas unidades. De algún modo los liderazgos sucesivos de Pura y Leonor en la dirección del departamento contribuyeron a evitar en buena medida las tendencias disgregadoras características de las estructuras de poder –como las universitarias– fuertemente politizadas. Los resultados de nuestra investigación mostraron que este éxito estuvo relacionado al menos con tres aspectos de los cuatro que exponemos a continuación:

- Una lectura adecuada de las características culturales que presentaban sus instituciones –en línea con los hallazgos de Kezar y Eckel (2002)–. Es posible que la mirada de las mujeres sobre las organizaciones y las relaciones sociales esté especialmente indicada para las transformaciones que las organizaciones universitarias están afrontando en estos momentos. En este sentido, en la elección de las mujeres para ocupar cargos directivos no estaría influyendo solamente el azar, la presión de éstas para conseguir mayor visibilidad o el deseo de acabar con las desigualdades del pasado, sino también la capacidad de las mujeres directivas para vincular los aspectos sociales al tratamiento de los problemas.
- La adopción de un estilo de liderazgo basado en la comunicación franca, abierta y directa, así como en la consideración de los problemas y necesidades personales de los miembros de la organización en equilibrio con los requerimientos de la tarea. Es probablemente la tendencia de las directivas a mantener un equilibrio entre la orientación hacia las personas y hacia las tareas lo que las conduce a buscar, como plantea Marshall (1990) formas creativas de resolución de conflictos, alejadas de las estrategias tradicionales basadas en los controles burocráticos. Esta tendencia proporciona al liderazgo de las directivas no tanto «visibilidad» como «profundidad». En efecto, comprobamos que el estilo de liderazgo de buena parte de las directivas analizadas no adoptaba una alta visibilidad en un sentido autoritario o bien ideológico. Los casos de Diana y Ofelia son paradigmáticos en este sentido. Por el contrario, se movía más bien en la *línea de sombra* de las relaciones interpersonales, multiplicando los contactos cara a cara y asumiendo las preocupaciones de los otros como la base de su actuación.
- La adopción también de un perfil resolutivo en la toma de decisiones. Si hay un perfil que predomina en los casos analizados es el que combina el despliegue de habilidades sociales, con la exigencia –incluida la autoexigencia– y la autoridad. Incluye una preocupación por la búsqueda del diálogo, del acuerdo y por limar las posiciones divergentes. Sin embargo, si el consenso no es posible no

se arredran en tomar decisiones aunque puedan levantar la oposición de un sector. Se trata por cierto de un perfil bastante «maternal», que conjuga la autoridad con la delicadeza o *diplomacia* en el manejo de las situaciones.

- Una fuerte implicación personal en las tareas, algo que aparece también en otros estudios (Loden, 1987). Algunos entrevistados en el caso de Diana destacan la especialmente intensa implicación de las directivas -Diana incluida- en las tareas y que incluso lo hacen «con más pasión», lo cual implica un mayor desgaste. Leonor por su parte ha llegado a percibir que «mi vida personal y mi vida profesional es la misma». Es decir, vive por y para el departamento con el que ha contraído un compromiso personal.

Todo esto nos lleva a sugerir que los modos de ejercer el poder característicos de las directivas de nuestro estudio se mostraron particularmente adecuados para llevar a cabo la clase de cambios que sus organizaciones necesitaban. Esto implica que si bien una parte de la explicación está en las características de las directivas, la otra se encuentra probablemente en las especiales características de las organizaciones educativas y, particularmente, de las de educación superior. Como han planteado Sánchez Moreno y López Yáñez (2003) la Universidad como organización posee una estructura débilmente acoplada. Esto quiere decir que su estructura no es el único patrón creador de ordenamiento organizativo, ni siquiera a veces el más poderoso. En estas organizaciones, las relaciones sociales y sus estructuras emergentes, informales y dinámicas, constituyen el principal factor de orden y también de desorden. Y se trata de estructuras por lo general más influyentes que las estructuras formales, en lo que se refiere a la creación de pautas colectivas de conducta y de pensamiento. Pues bien, quizás esas estructuras débilmente acopladas requieran, especialmente en estos momentos, formas «blandas» de ejercicio del poder, es decir, modos que no se apoyen predominantemente en la estructura formal, sino en el establecimiento de redes sociales y comunitarias, que impulsen la implicación y la creatividad del resto de los miembros de la organización.

## Limitaciones y prospectiva

La principal limitación de nuestro estudio se refiere a la profundidad que hubiera sido necesaria alcanzar en la fase de estudio de casos, que es la que ha sido descrita en este artículo. Aunque contábamos con la información obtenida en la fase descriptiva del

estudio, una mayor permanencia de los investigadores en las organizaciones analizadas habría permitido incrementar el rigor y el alcance de nuestras conclusiones. Sin embargo, el diseño de la investigación se ha mostrado prometedor para futuros estudios, especialmente la mencionada combinación de estrategias cuantitativas y cualitativas. Desde nuestro punto de vista, dichos estudios deberían en el futuro por una parte, establecer las diferencias que sobre los aspectos estudiados muestran los hombres que ejercen la dirección de este tipo de organizaciones; por otra, ampliar el análisis del liderazgo de las directivas a un rango más amplio de organizaciones universitarias, especialmente en lo que se refiere a sus culturas organizativas y estructuras de poder; y finalmente, realizar un seguimiento longitudinal de los procesos de transformación emprendidos por las directivas en el contexto de sus organizaciones.

## Referencias bibliográficas

- ACKER, S. (2005). Gender, leadership and change in faculties of education in three countries. En COLLARD, J. Y REYNOLDS, C., *Leadership, gender and culture in education. Male and female perspectives*. Maidenhead, UK: Open University Press.
- ALVESSON, M. Y BILLING, Y. (1992). Gender and Organization: Towards a Differentiated Understanding. *Organizations Studies*, 13 (12). 73-102.
- ALVESSON, M. Y DEETZ, S. (1996). Critical theory and postmodernism approaches to organizational studies. En S.T., CLLEGG, S.T., C. HARDI, C. Y W. R. NORD *Handbook of organizational studies* (p. 191-217). Londres: Sage, 191-217.
- BENSIMON, E.M. (1993). A feminist reinterpretation of president's definition of leadership. En J. GLAZER, J., E.M. BENSIMON, E.M. Y B.K. TOWNSEND. *Woman in higher education: a feminist perspective* (p. 465-474). Needham Heights, MA: Gin Press.
- BILLING, Y. Y ALVESSON, M. (2000). Questioning the notion of feminine leadership: a critical perspective on the gender labelling of leadership. *Gender, Work and Organization*, 7 (3), 44-157.
- BLACKMORE, J. (1999). *Troubling women: Feminism, leadership and educational change*. Buckingham: Open University Press.
- (2005). The emperor has no clothes'clothes: Professionalism, performativity and educational leadership in high-risk postmodern times. En J. COLLARD, J. Y C. REYNOLDS, C. *Leadership, gender and culture in education. Male and female perspectives*. Maidenhead, UK: Open University Press and McGraw Hill.

- CALÁS, M. B. Y SMIRCICH, L. (1993). Re-writing Gender into Organizational Theoring: Directions from Feminist Perspectives. EN J. GLAZER, E.M. BENSIMON Y B.K. TOWNSEND. *Woman in higher education: a feminist perspective* (p. 97-117). Needham Heights, MA: Gin Press.
- CHIA, R. (2003). Organization theory as a postmodern science. EN H. TSOUKAS, H. Y C. KNUDSEN. *The Oxford handbook of organization theory* (p. 113-140). Oxford: Oxford University Press.
- COLLARD, J. Y REYNOLDS, C. (2005). Conclusions and implications: Towards emergent theoretical perspectives on leadership, gender and culture. EN J. COLLARD, J. Y C. REYNOLDS. *Leadership, gender and culture in education. Male and female perspectives*. Maidenhead, UK: Open University Press and McGraw Hill.
- COURT, M. (2005). Negotiating and reconstructing gendered leadership discourses. EN J. COLLARD, J. Y C. REYNOLDS. *Leadership, gender and culture in education. Male and female perspective*. Maidenhead, UK: Open University Hill and McGraw Hill.
- ERKUT, S. (2001). *Inside Women´s Power: Learning from Leaders*. Boston, Center for Research on Women. Wellesley College, and Winds of Change Foundation.
- FLETCHER, J. K. (2001). *Disappearing Acts. Gender, Power and Relational Practice at Work*. Cambridge, MA: MIT Press.
- FULLAN, M. (2002). *Liderar en una cultura de cambio*. Barcelona: Octaedro.
- GHERARDI, S. (1995). *Gender, symbolism and organizacional cultures*. London: Sage.
- GROGAN, M.E. (1996) *Voices of women aspiring to the superintendency*. NY: State University of New York Press.
- HARGREAVES, A. (1996). *Profesorado, cultura y postmodernidad*. Madrid: Morata.
- KEZAR, A. Y ECKEL, P.D. (2002). The effect of institutional culture on change strategies in higher education: Universal principles/principles of culturally responsive concepts. *Journal of Higher Education*, 73 (4), 435-460.
- LEA, D. Y LEIBOWITZ, Z. (1992). *Adult Career Development. Concepts, Issues and Practice*. The National Career Development Association.
- LODEN, M. (1987). *Dirección femenina. Cómo triunfar en los negocios sin actuar como un hombre*. Barcelona: Hispano-Europea.
- MARSHALL, J. (1990). *Women Managers. Travellers in a male world*. Chichester: John Wiley & Sons.
- MILLS, A.J. (2002). Studying the gendering of organizational culture over time: Concerns, issues and strategies. *Gender, work and organization*, 9 (3), 286-307.
- NIDIFFER, J. (2001). New Leadership for a New Century. EN J. NIDIFFER, J. Y C.T. BASHAW. *Women administrators in higher education. Historical and contemporary perspectives*. New York: State University of N.Y. Press.

- PERREAULT, G. (1993). Contemporary feminist perspectives on women and higher education. En J. GLAZER, E.M.J., BENSIMON, E.M. Y B.K. TOWNSEND, *Woman in higher education: a feminist perspective* (p. 3-21). Needham Heights, MA: Gin Press.
- SÁNCHEZ MORENO, M. Y LÓPEZ YÁÑEZ, J. (2003). Condiciones organizativas de la enseñanza en la universidad. En C. MAYOR (coord.), *Enseñanza y Aprendizaje en la Educación Superior*. Barcelona: Octaedro.
- TIERNEY, W.G. Y BENSIMON, E.M. (1996). *Promotion and tenure. community and socialization in academe*. NY: State University of New York Press.
- WILLMOTT, H. (2003). Organizational theory as a critical science? Forms of analysis and 'new organizational forms'. En H. TSOUKAS, H. Y C. KNUDSEN, *The Oxford handbook of organization theory: metatheoretical perspectives*. Oxford: Oxford University Press, cap. 3.
- WOLF-WENDEL, L. Y WARD, K. (2003). Future prospects for women faculty. negotiating work and family. En B. ROPERS-HUILMAN, *Gendered futures in higher education. Critical perspective for change* (p. 111-136). NY: State University of New York Press.
- YANCEY MARTIN, P. Y COLLISON, D. (2002). Over the pond and across the water: Developing the field of gendered organizations. *Gender, Work and Organization*, 9 (3), 241-265.

**Dirección de contacto:** Julián López Yáñez. Universidad de Sevilla. Facultad de Ciencias de la Educación. Departamento de Didáctica y Organización. Calle Camilo José Cela, s/n. 41008, Sevilla, España. E-mail: lopezya@us.es



# Las escuelas eficaces: un estudio multinivel de factores explicativos del rendimiento escolar en el área de matemáticas

## The effective schools: a multilevel study of explanatory factors of the school performance in the area of mathematics

Covadonga Ruiz de Miguel

*Universidad Complutense de Madrid. Facultad de Educación. Departamento de Métodos de Investigación y Diagnóstico en Educación (MIDE). Madrid, España*

### Resumen

En este trabajo se presenta un estudio enmarcado en la línea de la eficacia escolar, línea clásica dentro de la investigación educativa iniciada con el Informe Coleman (1966), del que se cumplen ahora 40 años.

Basándonos en los factores que la literatura, española y americana, ha identificado como determinantes de la eficacia escolar, de entre los que cabe destacar factores relacionados con el propio alumno, el centro, el profesor y la familia, se ha elaborado un modelo explicativo del rendimiento de los alumnos que sintetiza los hallazgos teóricos citados y que incluye además el aspecto económico como predictor del rendimiento.

Para la elaboración de este trabajo se ha utilizado la metodología multinivel, metodología que propone una estructura de análisis dentro de la cual se pueden reconocer los distintos niveles en que se articulan los datos, estando cada subnivel representado por su propio modelo (Draper, 1995).

Para la validación del modelo se ha utilizado la base de datos del estudio PISA 2003, de la que se ha hecho una explotación parcial. La validación empírica ha permitido confirmar alguno de

los predictores, si bien no todos los incluidos en un principio, dando lugar a un modelo comprensivo que incluye los constructos recursos familiares; características personales del alumno; recursos de la escuela; contexto educativo del centro; clima escolar; e inversión económica

*Palabras clave:* calidad, modelo multinivel, eficacia escolar, Informe PISA, resultados escolares.

### **Abstract**

This paper presents a study is classified in the classical school effectiveness research agenda that begins with Coleman Report (1966) forty years ago.

Based on research literature, Spanish or American, an explanatory student achievement model has been built. It includes variables related to students, schools, teachers and student family. Among them we also include economical predictors, as amount of money devoted to education.

In order to do this survey, we use multilevel modeling, since this analysis allows differentiate hierarchical data structure, developing models for each data level (Draper, 1995). This analysis has been doing with PISA 2003 data base. Main results show a comprehensive explanatory achievement model that includes significant predictors as family resources, individual student features, school resources, educational context, school climate and economical investment.

*Key Words:* quality, multilevel model, school effectiveness, PISA report, school results.

## **Introducción**

En la investigación clásica sobre eficacia escolar se detectan algunas limitaciones como la imposibilidad de diferenciar efectos procedentes de los distintos agentes implicados en el sistema educativo, o la de establecer relaciones entre los diferentes niveles de agregación existentes. La metodología multinivel permite superar estas limitaciones, convirtiéndose en una herramienta metodológica de gran relevancia para la investigación educativa.

Este trabajo ha sido realizado con los datos del estudio PISA 2003, y a partir de ellos se ha elaborado un modelo de tres niveles (estudiantes, escuelas y países), explicativo del rendimiento en matemáticas de los alumnos de 15 años. La existencia de varianza dentro de cada nivel indica que los sujetos se diferencian entre ellos dentro de cada uno: los estudiantes difieren entre si, el rendimiento en matemáticas de las escuelas difiere de unas a otras, y entre los países ocurre lo mismo. Se incluyen en el

modelo variables relativas a cada uno de los niveles que la investigación ha señalado como factores relacionados con la eficacia escolar, medida a través del rendimiento de los alumnos en matemáticas.

El fin último de este trabajo es determinar la importancia relativa de los factores que afectan al rendimiento en este área, diferenciando estos factores entre los relativos al alumno, a la escuela y al país.

## Marco y Problema de Investigación: la eficacia escolar

La línea de investigación sobre eficacia escolar se puede considerar un *clásico* dentro de la investigación educativa. Su inicio lo marca el Informe Coleman (1966), centrado en la problemática de la desigualdad de resultados en educación y, desde su publicación, la investigación sobre este tema ha experimentado un cambio en cuanto a modelos, técnicas de análisis utilizadas, variables incluidas, instrumentos para la recogida de datos, etc.

Así, encontramos diferentes etapas que van desde la consideración exclusiva de variables de entrada para explicar el rendimiento de los alumnos (Coleman, 1966), pasando por otros más comprensivos que incluyen además aspectos de proceso (Brookover, 1979; Lezzote, Edmons y Ratner, 1974, etc.), hasta llegar a los que incluyen el contexto (Glasman y Biniaminov, 1981; Murphy, Hallinger & Mesa, 1985).

De todos estos trabajos surgen los cinco factores de Rutter et al. (1979), las listas de factores relacionados con la eficacia elaboradas por Reynolds (1985) en Inglaterra o Teddlie y Stringfield (1993) en EE. UU. En los últimos años, se ha intentado superar la elaboración de estas listas para proponer modelos comprensivos de eficacia, como los de Scheerens (1997) que pone el énfasis en aspectos relativos al funcionamiento de la escuela y Creemers (1994), centrado en la manera de trabajar en el aula.

Centrándonos en el proceso de enseñanza-aprendizaje, y a partir de otras investigaciones, en la búsqueda de factores y variables relacionadas con la eficacia escolar, nos encontramos con el *Informe de Seguimiento de la Educación para Todos en el Mundo*, elaborado por la UNESCO (2005), del que se deriva que «un proceso didáctico eficaz supone un entorno favorable, que consiste en *buenos profesores* (Babu y Mendro, 2003; Rivkin et al, 2002; citados en Gauthier y Dembele, 2004), *instalaciones adecuadas, suministro y distribución de materiales*

*didácticos* (Fullen y Clarke, 1994; Glewwe *et al*, 2000; Horsley, 2004), *autonomía*, así como una adecuada *estructura física* de la escuela y del aula».

El estudio sobre la eficacia escolar ha sido también trabajado en el contexto español. Como síntesis de las aportaciones realizadas, cabe señalar el modelo teórico de funcionamiento de los centros docentes elaborado por el *Centro de Investigación y Documentación Educativa* (CIDE) en 1995 con el fin de determinar su eficacia, tomando como referencia un modelo de calidad no centrado exclusivamente en el simple rendimiento académico de los alumnos. En este modelo, los factores más importantes hacen referencia a procesos de enseñanza, tanto en el ámbito *institucional* (en lo que se refiere al centro escolar como unidad institucional) como a nivel *didáctico* (trabajo llevado a cabo dentro cada aula).

En resumen, y a pesar de las diferencias metodológicas de las diferentes investigaciones, casi todas coinciden en afirmar cuáles son los factores determinantes de la eficacia escolar en España (Murillo, 2000): una dirección adecuada, el liderazgo pedagógico (Rodríguez, 1991), el buen funcionamiento del equipo directivo (García, 1991) y la valoración positiva por parte de la comunidad escolar hacia el trabajo de los directivos (Muñoz Repiso *et al*, 1995; Castejón, 1996); el *clima escolar* (Muñoz Repiso, *et al*, 1995; Fuentes, 1986; Rodríguez, 1991 y Castejón, 1996); el *trabajo innovador y en equipo de los profesores* (Rodríguez, 1991; Muñoz Repiso *et al*, 1995; Castejón, 1996); la *participación e implicación de las familias* (García, 1991; Rodríguez, 1991).

Todas las investigaciones realizadas coinciden en identificar los mismos factores como responsables de la eficacia escolar. Así pues, el modelo que se propone reúne una serie de constructos que la literatura ha identificado en los diferentes estudios realizados como determinantes de la eficacia escolar. Concretamente estamos hablando de los constructos: *recursos familiares, características personales del alumno, recursos de la escuela, contexto educativo del centro y clima escolar*. El modelo incluye variables de tipo económico, fruto de la reflexión acerca de su relación con la eficacia escolar, que parece sugerir que la obtención de buenos resultados por parte de los alumnos se haga sin malgastar esfuerzo ni recursos, optimizando todo aquello de que se disponga (Hanushek, 1992).

En este sentido, recientemente se ha incluido el estudio de lo que cada país invierte en educación por la incidencia que pudiera tener esta variable sobre los resultados escolares. A este respecto, se propone que los países inviertan en educación, por lo menos, el 6% del Producto Interior Bruto (PIB) (UNESCO, 1996), si bien esto no sería garantía de calidad, si que puede considerarse una medida de valor político

ya que, alcanzando este objetivo, elevaría en muchos países el índice de recursos disponibles (UNESCO, 2005).

Diferentes trabajos han intentado determinar si existe una relación causal entre el gasto en educación y los resultados globales. El conjunto más sólido de resultados parece confirmar la importancia de los recursos escolares para las puntuaciones en los tests (Lee y Barro, 2001). Otros estudios demuestran que el gasto en la enseñanza primaria (no así el gasto total en educación) influye en el rendimiento hasta los grados 4º y 5º (Gupta, Verhoeven y Tiongson, 1999; McMahon, 1999). Otros resultados, sin embargo, son muy diferentes. Hanushek y Kimko (2000) por ejemplo, observan que las variaciones en los recursos escolares no tenían repercusiones muy significativas en los resultados de los tests escolares. Según sus estimaciones, los efectos de diferentes tipos de recursos son estadísticamente irrelevantes o, más frecuentemente, estadísticamente significativos, pero inversos a los esperados.

Para la elaboración de este trabajo partimos de la consideración de escuela eficaz definida como *aquella en la que los alumnos consiguen un rendimiento óptimo* (Muñoz Repiso, 1997). De este modo podría decirse que la eficacia y la mejora escolar tienen sentido en función del logro de la calidad educativa, cifrada en la consecución de unos resultados amplios, que deben abarcar todas las facetas de la formación humana y no excluir a ninguna persona: al medir los resultados es preciso tener en cuenta el contexto, las condiciones de los alumnos, su nivel de partida; y desde esta perspectiva los procesos (recursos, medios, etc.) más eficaces serán los que consiguen mejores resultados. La escuela eficaz es, por tanto, la que consigue los fines propuestos con los medios adecuados y la escuela de calidad es la que, además, se propone fines social y humanamente relevantes (Muñoz Repiso, 1997).

## Planteamiento del modelo de investigación

El objetivo general de este trabajo es determinar la importancia relativa de los factores que afectan al rendimiento en matemáticas, en cada uno de los niveles de agregación de los datos, el primer nivel individual, el de la escuela y el de el país.

De los diferentes factores que la literatura identifica como relacionados con la eficacia escolar, para la elaboración del modelo que se propone se han seleccionado los constructos latentes que se presentan en la Tabla I junto con sus variables observadas.

## Diseño y metodología

Desde el punto de vista metodológico, los estudios sobre eficacia escolar han utilizado tradicionalmente el análisis de regresión (García, 1991; Rodríguez Gómez, 1991, Castejón, 1996). Las limitaciones de esta metodología (Castejón, 1996) abrieron la puerta a la aparición de los Modelos Multinivel, alternativa metodológica que se ha mostrado como la más adecuada para la investigación sobre eficacia escolar (Murillo, 2000). Estos modelos ofrecen la posibilidad de poder recoger la estructura anidada de los datos educativos en sus niveles de alumno, centro y país, así como poder distinguir con mayor precisión los efectos debidos a cada uno de esos niveles. De la misma forma, y esto es conceptualmente muy destacable, el modelo incluye en la posibilidad de reconocer la equidad del centro, elemento fundamental para una correcta concepción de los términos *eficacia escolar* y *calidad de la educación* (Gaviria y Castro, 2004).

## Población y muestra

Para la realización de este trabajo se ha utilizado la base de datos procedente del estudio PISA 2003<sup>1</sup>. La población objetivo en PISA son los alumnos de 15 años de los países de la OCDE (más alguno otro que no pertenece a ella). Los datos recogidos en PISA se organizan jerárquicamente, en tres niveles: *alumnos*, *escuelas* y *países*. En principio, el estudio PISA contempla la existencia de 41 países, 10274 escuelas y 276165 estudiantes.

El programa que se ha utilizado para el análisis de los datos ha sido MLwin 2.0. La exportación de datos desde SPSS ha producido una pérdida de casos al eliminarse todos aquellos en los que aparecen valores perdidos en alguna variable (*listwise*). La muestra que será utilizada en este trabajo queda configurada por 35 países y 100325 estudiantes.

La variable de respuesta utilizada como variable dependiente en este trabajo se ha elaborado a partir de los cinco los valores plausibles *en matemáticas*, que el estudio PISA asigna para cada estudiante en esta materia<sup>2</sup>.

## Procedimiento de análisis

Las variables que se introducirán como predictores en cada nivel del modelo son las siguientes (Tabla D):

<sup>1)</sup> Los datos están disponibles en la página [http://pisaweb.acer.edu.au/oecd\\_2003/oecd\\_pisa\\_data.html](http://pisaweb.acer.edu.au/oecd_2003/oecd_pisa_data.html).

<sup>2)</sup> El área principal en PISA 2003 fue la competencia en matemáticas, mientras que la lectura, las ciencias y la solución de problemas fueron áreas secundarias. Los datos cognitivos se escalaron con el modelo de Rasch y el rendimiento de los alumnos se expresó mediante la asignación de *valores plausibles*.

TABLA I. Relación de variables independientes agrupadas por niveles<sup>3</sup>

Nivel 1. alumno	
Recursos familiares	Recursos educativos en el hogar (Hedres) Posesiones en el hogar (Homepos) Recursos culturales en hogar (Cultpos) Nivel socioeconómico familiar (Escs)
Características del alumno	<b>Personales</b> Sexo (Sexo) Tiempo dedicado a deberes de matemáticas (Rmhhmk)
	<b>Actitudinales</b> Aspiraciones educativas de los estudiantes (Sised) Actitudes de los estudiantes hacia la escuela (Atschl) Sentimiento de pertenencia al centro (Belong) Autoeficacia en matemáticas (Matheff) Autoconcepto en matemáticas (Scmat)
	<b>Cognitivas</b> Control (estrategia de aprendizaje) (Cstrat) Elaboración (estrategia de aprendizaje) (Elab) Interés en matemáticas (Intmat)
	<b>Afectivas</b> Disciplina en el aula (Discim) Relaciones entre alumnos y profesores (Sturel) Apoyo del profesor (Techsup)
Nivel 2. escuela	
<b>Capital humano de la Escuela</b>	Proporción de profesores de matemáticas con certificación 5a (Propma5a)
<b>Recursos de la escuela</b>	Calidad de los recursos físicos de la escuela (Scmatbui) Calidad de los recursos educativos de la escuela (Scmatedu)
<b>Contexto educativo del centro</b>	Ratio alumno/profesor de matemáticas (Smratio) Frecuencia de evaluaciones (Asses) Agrupamiento de los alumnos (Abgroup) Actividades extraordinarias de matemáticas (Mactiv) Cursos extraordinarios de refuerzo (Excouse) Autonomía de recursos (Autres) Autonomía curricular (Autcurr) Autonomía a de la escuela (Schauton) Participación de los profesores (Tchparti)
<b>Clima escolar</b>	Relaciones entre profesores y alumnos (Mstrel) Entusiasmo y compromiso de los profesores (Tcmorale) Entusiasmo y compromiso de los estudiantes (Stmorale) Conducta del profesor (Teacbeha) Conducta de los alumnos (Stbeha) Consenso entre los profesores (Tchcons)
Nivel 3. país <sup>4</sup>	
<b>Económicas</b>	PIB del país <sup>5</sup>
<b>INVERSIONES EN EDUCACIÓN</b>	INVERSIONES EN EDUCACIÓN % del PIB que se invierte en Educación Gasto en educación por alumno % del PIB destinado a Educación Secundaria

<sup>3</sup> En el apéndice 9 del Manual para usuarios de SPSS publicado por la OCDE (<http://www.oecd.org/dataoecd/35/14/35003010.pdf>) puede consultarse la definición detallada de cada una de las variables y cómo PISA las ha construido.

<sup>4</sup> Fuente: OCDE (2004). *Education at glance*. París: ECD. [www.oecd.org/edu/eag2004](http://www.oecd.org/edu/eag2004)

<sup>5</sup> Todas las cantidades, en dólares americanos, están equiparadas en *Purchasing Power Parity* (ppp). *Purchasing Power Parity*: Dólar Internacional.

Las variables *tiempo dedicado a deberes* y *% de PIB que cada país invierte en Educación* han sido centradas respecto de la media de la muestra con el fin de facilitar la interpretación de resultados. Para ello se ha calculado la diferencia entre el valor original en PISA 2003 y la media muestral, dando lugar a un nuevo valor más fácilmente interpretable.

## Análisis y resultados

El cálculo de los parámetros se ha realizado siguiendo la *metodología de valores plausibles* (Wu y Adams, 2002), estimándose para cada uno de los cinco valores el modelo y calculando los promedios (Wu y Adams, 2000), lo que dará lugar a la estimación tanto de los parámetros como de sus errores típicos. Estos parámetros promediados son los que se presentan en este informe de resultados.

Antes de comenzar con el proceso de modelización, se ha calculado la autocorrelación o correlación intraclase, que expresa el grado de homogeneidad de los contextos, ignorada habitualmente por los modelos estadísticos clásicos. Por motivos de espacio no presenta más que los resultados, siendo 0,2252 la autocorrelación de alumnos dentro de escuelas, 0,242 la de los alumnos dentro de países, y 0,518 la de escuelas dentro de países.

Estos tres valores muestran homogeneidad de las unidades dentro de cada nivel, lo que justifica el empleo de los modelos multinivel, que consideran en su proceso de estimación la presencia de correlación intraclase.

A continuación se presenta parte del proceso de modelización: el modelo nulo y el de tercer nivel, sobre el que se hace la interpretación final.

El modelo nulo (Tabla ID), se establece como línea base para la evaluación comparada de los modelos. Aporta la partición básica de variabilidad de los datos en dos o más niveles, y permite evaluar la adecuación del uso de modelos multinivel (Gaviria y Castro, 2004).

En este modelo se calculan cuatro parámetros: la constante en la parte fija y las varianzas de los residuos en los tres niveles en la parte aleatoria. La variable de respuesta en el modelo nulo es el *rendimiento en matemáticas de los alumnos de Secundaria*.

El criterio para decidir si un parámetro es o no significativo es que, para  $\alpha=0,05$ , el cociente entre la estimación del parámetro y su error típico sea superior a 1,96 (Gaviria y Castro, 2004). Este es el criterio que seguiremos de aquí en adelante.

TABLA II. Estimación del modelo nulo

Parte Fija		
Parámetro	Estimación	S. Error
Constante	509,1566	8,327
Parte Aleatoria		
Nivel 3		
Parámetro	Estimación	S. Error
Constante	2384,7632	582,5552
Nivel 2		
Parámetro	Estimación	S. Error
Constante	2219,6512	51,7268
Nivel 1		
Parámetro	Estimación	S. Error
Constante	5248,454	24,0198
$2^{\circ}$ loglikelihood (IGLS Deviance) $X^2_{(15)} = 1154766,2$ (100325 of 100325 cases in use)		

Como se aprecia en la Tabla II, la media de rendimiento en matemáticas para todos los sujetos es de 509,1566 puntos. Los sujetos difieren entre ellos en rendimiento: la estimación de la varianza en nivel estudiantes es de 5248,454, y su error típico es de 24,0198. Podemos observar en la Tabla II que las escuelas también difieren en su rendimiento medio ( $2219,6512 / 51,7268 = 42,91$ ), y que los países también presentan rendimientos medios estadísticamente distintos ( $2384,7632 / 582,5552 = 4,09$ ).

La significación de estos parámetros indica una variación sin explicar entre los rendimientos de los alumnos, entre los rendimientos de las escuelas, y también entre los rendimientos entre los países. La razón de verosimilitud tiene un valor de 1154766,2 para un modelo con cuatro parámetros. Al finalizar el proceso de modelización, este valor, será comparado con el ofrecido al estimar el modelo definitivo, para evaluar su ajuste.

Comprobada la significatividad del modelo nulo, trataremos de explicar la mayor cantidad de varianza posible, expandiendo el modelo a través de la incorporación de predictores de primer, segundo y tercer nivel en la parte fija y aleatoria.

El proceso de modelización se hace por partes, en primer lugar se introducen las variables de primer nivel incluidas en la Tabla II, para posteriormente proceder con las de nivel 2 y 3. Como resultado de la estimación de los sucesivos modelos se han desestimado variables por no resultar sus parámetros significativos. Las interacciones entre los predictores de nivel dos no han resultado significativas, como tampoco lo han sido las interacciones de variables interniveles, por lo que no se considera ninguna.

Para la estimación del modelo definitivo, se han introducido variables predictoras de nivel 3. Las variables relativas al PIB del país y a la inversión anual que hacen los diferentes países en materia de educación diferenciando por niveles educativos (en

dólares americanos equiparados a través del ppp<sup>6)</sup>, se han desestimado al no resultar sus parámetros significativos. Sin embargo, la variable relativa al % de PIB que cada país invierte en educación, equiparado igualmente en *dólares ppp*, resultó significativo. Esta variable fue centrada respecto de la media para facilitar la interpretación del parámetro obtenido. La interacción entre las variables de tercer nivel no resultó significativa, y no se consideró oportuno calcular interacciones entre éstas variables y otras de otros niveles al no encontrar argumentos teóricos que lo justificasen.

El resultado de la estimación del modelo definitivo, con predictores de nivel 3, se presenta en la Tabla III.

TABLA III. Modelo con variables explicativas en el nivel 3 (modelo definitivo)

Parte fija		
Parámetro	Estimación	S. Error
Constante	448,984	7,8494
Sexo	8,4796	0,4682
Escs	10,2238	0,3358
sisced	12,8338	0,2324
scmat	16,5378	0,269
matheff	22,9404	0,268
rmhmkw (c)	-25,5266	0,7412
disclim	4,348	0,2334
Escs*sex	-2,5818	0,420
abgroup	-1,8532	0,691
schauton	1,6054	0,556
propma5a	20,0038	3,8808
scmatedu	5,349	0,581
scmatbui	-2,6602	0,589
stmorale	4,1118	0,452
tcmorale	2,002	(0,395)
studbeha	2,8562	(0,355)
asses	-3,0144	(0,837)
mstrel	-73,178	(12,551)
excource	2,1496	(0,808)
%GPD (c)	11,9558	(5,9)
Parte aleatoria		
Nivel 3		
Parámetro	Estimación	S. Error
Constante	1750,8262	438,5526
%GPD/%GPD	300,6118	114,9618
Nivel 2		
Parámetro	Estimación	S. Error
Constante	1348,9062	438,5526
Sex/sex	116,3168	16,4934

<sup>6)</sup> Purchasing Power Parity. Dólar Internacional.

escs/escs	27,7286	5,7814
sisced/sisced	39,8684	4,1266
scmat/scmat	43,3826	5,0532
rmhmwk(c)/rmhmwk(c)	321,2402	47,8656
disclaim/disclim	31,9508	4,6188
<b>Nivel I</b>		
<b>Parámetro</b>	<b>Estimación</b>	<b>S. Error</b>
Constante	3521,342	17,7532
$-2^{\circ}\log\text{likelihood (IGLS Deviance)} \chi^2_{P(1,5)} = 1116386,4 (100325 \text{ of } 100325 \text{ cases in use})$		

La estructura de la salida es la misma que la del modelo nulo (parte fija, parte aleatoria y razón de verosimilitud). Sin embargo, si la comparamos con aquel, la incorporación de predictores añade términos a la parte fija y aleatoria del modelo.

La interpretación que hacemos del modelo definitivo, tal y como se aprecia en la Tabla III, es la siguiente: El valor de la constante es 448,984 puntos, que representa la media esperada en rendimiento en matemáticas para las *chicas*, con un *nivel socioeconómico familiar* igual a la media de la muestra, sin *aspiraciones educativas*, con un *autoconcepto en matemáticas* y una *autopercepción de eficacia en matemáticas* iguales a la media de la muestra, con un *tiempo proporcional dedicado deberes de matemáticas* igual a la media de la muestra, y con una percepción neutra del *clima de aula*, que asisten a escuelas en las que *no hay agrupamiento por habilidad*, con niveles medios de *autonomía escolar*, con un *número de profesores de matemáticas con certificación 5ª* igual a la media de la muestra (0,71187), con niveles de *calidad de recursos, tanto físicos (infraestructuras) como educativos*, iguales a la media de la muestra, con niveles medios de *moral y compromiso de profesores y estudiantes, de profesores y estudiantes*, y de conducta de los estudiantes, en los que se realizan hasta 20 *evaluaciones*, en las que el índice de  *Pobres relaciones entre estudiantes y profesores* es igual a la media de la muestra, y en la que no se ofrece ningún tipo de *actividades extra*, y que pertenecen a un país que *invierte en educación una parte del PIB* que coincide con la media de la muestra (5,16%).

La media de rendimiento en matemáticas esperada (448,9408) sería de 8,4796 puntos más si se tratase de un chico. Por cada punto que aumentase el *nivel socioeconómico familiar*, la media de rendimiento esperado aumentaría 10,2238 puntos más, y otros 12,8338 por cada nivel educativo que aumenten las *expectativas educativas del alumno*. Por cada punto de aumento en el *autoconcepto* en matemáticas y la *autopercepción de la eficacia* en esta área, el rendimiento esperado aumentaría 16,5378 y 22,9404 respectivamente. Por cada punto que aumente la proporción de *tiempo de deberes* que dedica a deberes de matemáticas, el rendimiento esperado disminuiría

-25,5266 puntos. Se produciría un aumento de 4,348 puntos por cada punto de aumento en la percepción de un clima de clase positivo. La existencia de una interacción entre las variables *sexo* y *nivel socioeconómico* indica que el rendimiento de los alumnos disminuiría en -2,5818 puntos en el caso de los chicos cuyo nivel socioeconómico familiar esté por debajo de la media, esto hace suponer que los chicos son más sensibles a un bajo nivel socioeconómico que las chicas, cuando siempre se ha pensado que el efecto tendría la dirección contraria. La explicación la podríamos encontrar en que las chicas tienen una mayor capacidad de adaptación al trabajo escolar que los chicos. Puede ocurrir que la permanencia de las chicas en la escuela esté más amenazada cuando el nivel socioeconómico es bajo, pero la calidad del trabajo escolar de éstas es mayor que la de los chicos cuando las primeras permanecen en la escuela, esto queda reflejado en una menor disminución del rendimiento en matemáticas cuando disminuye el nivel socioeconómico.

La interpretación de los parámetros de nivel escuela es la siguiente: a media que aumenta el *agrupamiento por habilidad* de los alumnos, el rendimiento esperado disminuye -1,8532 puntos. El rendimiento esperado aumenta otros 1,6054 puntos por cada punto que aumenta la *autonomía escolar*; y 20,0038 puntos más por cada punto de aumento de proporción de *profesores con de matemáticas con certificación pedagógica 5ª*.

Por cada punto que aumenta la *calidad de los recursos educativos* del centro aumenta el rendimiento esperado de los alumnos 5,349 en puntos, mientras que al aumento en la *calidad de los recursos físicos* le correspondería una disminución de -2,6602 puntos.

Al aumento de cada punto en el nivel de *entusiasmo y compromiso de los estudiantes y profesores*, le corresponde un aumento en el rendimiento de 4,1118 y 2,002 puntos respectivamente. A cada punto de aumento de *buenas conductas de los alumnos* le corresponde un aumento de 2,8562 puntos.

A medida que las *evaluaciones* se hacen más frecuentes, el rendimiento disminuye -3,0144 puntos. La existencia de *pobres relaciones entre estudiantes y profesores* hace que el rendimiento esperado disminuya -73,178 puntos. Y, por último, la existencia de actividades extraordinarias hace que el rendimiento esperado aumente 2,1496 puntos más.

La inclusión del predictor de nivel 3 (país) nos permite afirmar que por cada punto porcentual que aumenta la *inversión del país en educación*, el rendimiento de los alumnos se vería beneficiado con un aumento de 11,9558 puntos sobre la media estimada.

Con relación a la parte aleatoria del modelo, habría que señalar, en primer lugar, el notable descenso de la varianza que queda por explicar entre el modelo nulo y el modelo final.

La parte aleatoria del modelo muestra algunas variaciones respecto del modelo nulo. La variación residual de las constantes en el tercer nivel muestra que, a pesar de la inclusión como predictor del % del PIB destinado a educación, todavía se observan diferencias en el rendimiento medio de matemáticas entre países. Los países también se diferencian en la relación que mantiene el rendimiento en matemáticas con la proporción de profesores con certificación pedagógica 5a (ver estructura de variación del nivel 3 en la Tabla III).

Como era de esperar, las escuelas siguen diferenciándose en su rendimiento medio en matemáticas. El rendimiento en matemáticas diferenciado entre escuelas está vinculado a las variables sexo, nivel socioeconómico familiar, aspiraciones educativas de los estudiantes, autoconcepto en matemáticas, tiempo dedicado a deberes de matemáticas y disciplina en el aula, tal y como se observa en la estructura de variación del 2º nivel (Tabla III).

En suma, vemos que se produce una disminución de la varianza en todos los niveles. Comparando la varianza que queda sin explicar en el modelo nulo y en el definitivo, apreciamos una disminución de esta en los tres niveles. En el nivel tres pasa de 2384,7632 a 2051,438 (para llegar a este valor basta con sumar las varianzas asociadas a cada predictor de nivel 3); en el nivel dos pasa de 2219,6512 a 1929,3936; y en el nivel uno, la varianza sin explicar pasa de 5248,454 en el modelo nulo, a 3521,342 en el definitivo.

Para la interpretación del modelo, además de analizar la significación de los coeficientes, debemos analizar su ajuste global. Debemos considerar si, además de ser significativa la aportación de los predictores, el modelo propuesto aporta información significativa comparándolo con el modelo nulo, presentado en la Tabla II. Las decisiones de mantener o eliminar un predictor y mantener o eliminar un modelo se hayan íntimamente ligadas, ya que no puede haber modelos significativos que incluyan predictores que no lo sean (Gaviria y Castro, 2004).

Para llevar a cabo la comparación, se hace uso de la razón de verosimilitud. El valor de  $-2 \cdot \log \text{likelihood}$  (IGLS) para este último modelo es de 1116386,4. Si lo comparamos con el valor obtenido en la estimación del modelo nulo, la diferencia entre ellos es de 38379,8. Para comprobar el ajuste del modelo, calculamos la significatividad de este valor con 29 grados de libertad (33 parámetros estimados en el modelo definitivo menos 4 parámetros estimados en el modelo nulo). Al obtener un valor de  $p=0,000$ , menor que el nivel de confianza con el que trabajamos ( $\alpha=0,05$ ), se estima que el modelo definitivo explica mayor varianza que el modelo nulo.

## Discusión de resultados y conclusiones

De la validación del modelo se puede extraer como conclusión general que contribuye al estudio de la eficacia escolar, al haber identificado una serie de factores, con respaldo teórico, que tomados conjuntamente, influyen significativamente en el rendimiento en el área de matemáticas de los alumnos de Secundaria.

Las conclusiones más específicas, que pasamos a comentar y discutir de forma más detallada, las hemos agrupado en torno a los factores que incluía el modelo propuesto.

En primer lugar el sexo del alumno sigue siendo un importante predictor de su rendimiento, al menos en el área de matemáticas. Confirmando los trabajos de algunos autores (Roa, 2006) del modelo se deriva que el rendimiento en matemáticas de los alumnos es superior en 8,5 puntos respecto del obtenido por las alumnas.

Algunos trabajos consultados revelaban que el tiempo *dedicado a los deberes* (CIDE, 1995) y las *aspiraciones educativas* (Martín Rodríguez, 1985) parecían tener relación con la eficacia de la escuela. En este trabajo se confirma este punto, si bien en su interpretación hay que hacer algunas matizaciones. Respecto del tiempo que el alumno dedica a realizar tareas escolares relacionadas con las matemáticas, ocurre un fenómeno, cuando menos curioso. El parámetro asociado a la variable tiene un valor de -25,5266, lo que en un principio podría ser interpretado como que «cuanto más tiempo se dedica a los deberes de matemáticas, menor es el rendimiento». Sin embargo, hay que tener en cuenta cómo se ha introducido esta variable en el modelo. En primer lugar, la variable ha sido centrada respecto de la media de la muestra, 0,48, este valor indica que, por término medio, los alumnos dedican casi la mitad del tiempo total de deberes a los deberes de matemáticas. El valor del parámetro indicaría que, por cada punto que aumente el porcentaje de tiempo que el alumno dedica a matemáticas, su rendimiento disminuiría 25,5 puntos. Esto podría estar sucediendo a los alumnos que tienen más dificultades con las matemáticas, que tienen que dedicarles un tiempo proporcionalmente mayor de deberes, y no siempre con buenos resultados. Este argumento se apoyaría al observar la correlación entre las variables *tiempo que dedica el alumno a deberes de matemáticas y rendimiento en matemáticas*, que ofrece un valor de 0,193 ( $p = 0,000$ ). Si bien la correlación es baja, es significativa. Por tanto, parece que dedicar hasta el 48% del tiempo total de deberes a deberes de matemáticas, sería beneficioso para el rendimiento. Sin embargo, cabe señalar que se trata de un efecto aparente (*prima facie*), al que no se puede dar una interpretación causal, ya que carecemos de la información suficiente para afirmar esta causalidad rotundamente.

Respecto de las *aspiraciones educativas del alumno*, también parecen ser un importante condicionante para su rendimiento. Los niveles que presenta esta variable son: 0=Educación Infantil; 1=Educación Primaria; 2= Ed. Secundaria (Inf.); 3= Educación Secundaria Superior. Bachillerato; 4=Estudios Universitarios; 5=Estudios tercer ciclo. Sobre la media de rendimiento estimada, que sería para un alumno sin aspiraciones educativas (valor 0), por cada punto que aumentase el nivel de expectativas, el rendimiento se vería considerablemente incrementado (casi 13 puntos). Cabe resaltar la importancia de este dato, ya que estamos hablando de una variable sobre la que es muy posible influir por medio de campañas de información o de otro tipo, a diferencia del nivel socioeconómico, variable sobre la que el sistema no tiene capacidad de acción.

Otros factores actitudinales que parecen influir en el rendimiento de los alumnos, confirmando otros trabajos (Rodríguez Espinar, 1979; Gimeno, 1976, Gonzalez y Tourón, 1992), son la *autopercepción de la eficacia en el desarrollo de actividades matemáticas*, y el *autoconcepto en esta materia*, que hacen que el rendimiento aumente 16,5 y casi 23 puntos respectivamente.

También variables afectivas, relacionadas con la *disciplina que el alumno percibe en el aula* parecen tener influencia sobre el rendimiento, en la línea de otros estudios anteriores. Parece que el aumento de la percepción de un buen clima en este sentido, hace que el rendimiento aumente 4,5 puntos, confirmando los trabajos de Bosch (1981) entre otros.

Respecto de los *recursos familiares*, el *nivel socioeconómico de la familia* confirma investigaciones pasadas (Pérez Serrano, 1981), que ponen de manifiesto la influencia que ejerce sobre el rendimiento, quizás por las oportunidades en cuanto a actividades y materiales educativos que esto implica. Según los resultados obtenidos, el alejarse de la media del nivel socioeconómico de la muestra supone un incremento (o disminución, en función de la dirección en la que aleje) de 8,5 puntos sobre la media esperada de rendimiento. Se constata en el modelo además una interacción entre esta variable y el sexo. Es decir, el rendimiento disminuiría -2,5 puntos adicionales en el caso de ser un chico y pertenecer a una familia con nivel socioeconómico inferior a la media.

Si bien estos aspectos pueden considerarse de *entrada*, al venir impuestos desde fuera y quedar poco margen de manipulación, de cara a la eficacia escolar que se busca, y en vista de su relación con el rendimiento, se podría trabajar en alguno de ellos. Está claro que el sexo de los alumnos, o el nivel socioeconómico de la familia poco margen de manipulación ofrecen, pero si que se puede trabajar con las expectativas educativas del alumno (a través de un buen sistema de incentiación e información sobre estudios

y salidas profesionales), la autoeficacia y autoconcepto en matemáticas (trabajando primero el área de las matemáticas y promoviendo un cambio de aptitudes respecto a ellas, lo que repercutiría en el tiempo necesario para hacer deberes...).

El *clima escolar*, por su parte, se revela en este modelo como un factor influyente sobre la eficacia escolar, una vez parcializado el efecto del nivel socioeconómico, como era de esperar a partir de otros trabajos consultados. Muñoz Repiso, et al. (1995), Rodríguez (1991) y Castejón (1996) entre otros.

Parece que en este sentido, la ausencia de «pobres relaciones entre alumnos y profesores», aumenta de forma considerable el rendimiento de los alumnos. Para elaborar este índice, se preguntaba a los directores de los centros acerca de *cómo se llevan los profesores con los alumnos, si se interesan por ellos; si les escuchan; si les dan ayuda extra; si les tratan correctamente, etc.* En la medida que se dan respuestas satisfactorias, se considera la no existencia de *pobres relaciones* entre ambos. Pues bien, el modelo pone de manifiesto que, en la medida que aumenta la existencia de malas relaciones, el rendimiento disminuye nada menos que 73 puntos. Parece que este aspecto es especialmente relevante, por lo que habrá que trabajar para lograr en el centro un adecuado nivel de relaciones entre alumnos y profesores, de cara a favorecer la eficacia de la escuela.

En la misma línea, el *entusiasmo y compromiso*, tanto de profesores como de alumnos, favorece el rendimiento en matemáticas de estos últimos, por lo que puede considerarse un factor más de eficacia escolar. Concretamente, el rendimiento aumenta unos 4 puntos cuando se perciben niveles de entusiasmo y compromiso por parte de los alumnos por encima de la media, y en torno a 2 puntos más en el caso de niveles de entusiasmo y compromiso de los profesores por encima de la media.

Respecto de la conducta del alumno, cabe resaltar que se produce un aumento de casi 3 puntos en la media estimada de rendimiento cuando la conducta del alumno aumenta respecto de la media (esta variable hace referencia a la ausencia de *absentismo escolar; interrupciones de las clases por parte de los alumnos, falta de respeto hacia los profesores, problemas de alcohol y/o drogas en los alumnos, o conductas de bullying en las aulas*).

Podría decirse que estas tres variables están muy relacionadas, de hecho, las correlaciones calculadas entre todas ellas resultan significativas trabajando con un  $\alpha$  de 0,05.

Estos aspectos, unidos a la variable anteriormente comentada, la *existencia de pobres relaciones entre alumnos y profesores*, darían idea del *clima* del centro, y en nuestro modelo se revelan como factores de eficacia escolar. Estos aspectos son igualmente

susceptibles de modificación, en el sentido de que se puede trabajar con alumnos y profesores para lograr altos niveles de entusiasmo y compromiso por parte de ambos, y altos niveles de buena conducta por parte de los estudiantes, lo que haría disminuir la existencia de *pobres relaciones entre ambos colectivos*.

Respecto del profesorado, la figura del docente ejerce una notable influencia sobre el rendimiento de los alumnos (Grahay, 2000). Según el modelo, parece que la proporción de profesores de matemáticas con certificación pedagógica 5ª (que correspondería al *Curso de Aptitud Pedagógica* en España) que haya en el centro es importante para el rendimiento de los alumnos. El resultado parece evidente: cuanto mayor es el porcentaje de profesores con esta certificación, mayor probabilidad hay de que el alumno tenga un profesor que conozca técnicas pedagógicas (estrategias didácticas y de evaluación, teorías de aprendizaje, etc.) adecuadas para contribuir a un alto desempeño por parte de los alumnos, lo que confirma algunos trabajos, como el de Walberg (1991). La media de profesores con esta certificación en los centros de la muestra es del 72%, y por cada punto que se separa de esta media, aumenta (o disminuye) el rendimiento esperado del alumno 20 puntos.

Los *recursos de la escuela* se muestran también como decisivos para la eficacia de la misma. Respecto de la calidad de los recursos, tanto educativos como físicos, se han encontrado resultados sorprendentes. Respecto de los primeros, la evidencia empírica habla de la relación existente entre unos materiales educativos de calidad en el centro y un alto desempeño por parte de los estudiantes (Fullen y Clarke, 1994; Walberg, 1991). En el modelo esto parece quedar probado, al encontrar un incremento de 5,3 puntos en la media de rendimiento en matemáticas estimado por cada punto que aumento la calidad de los recursos educativos. Lo sorprendente lo encontramos al valorar el parámetro asociado a la calidad de los recursos físicos del centro (edificio y terrenos; sistemas de calefacción, aire acondicionado y electricidad; aulas; etc.), que en estudios anteriores se asocia con la eficacia escolar (Edmonds, 1979, entre otros). En nuestro trabajo hemos encontrado un valor desconcertante, pues parece que un aumento en la calidad de los recursos físicos de la escuela, produciría una bajada en el rendimiento de 2,6 puntos.

Recordemos que esta variable es un índice que PISA construye a partir de las respuestas que los directores de las escuelas dan a las preguntas: *¿En su escuela la capacidad de proporcionar instrucción se ve afectada por la ausencia o falta de adecuación del edificio o terrenos, sistemas eléctricos, de calefacción o aire acondicionado, o de espacios de instrucción (como aulas)?* Las respuestas se valoran de 1 (en absoluto) a 4 (Mucho) y las medias en estas variables están en torno a la

categoría 2 (muy poco), lo que implicaría que, por término medio, los directores consideran que la calidad de los centros que dirigen es buena, lo que podría explicar el signo del parámetro encontrado.

En la investigación revisada para la elaboración del marco teórico del modelo, encontrábamos un indicador de los recursos escolares, la *proporción de alumnos por docente*, con incidencia positiva sobre los resultados educativos. En nuestro modelo esta relación no se confirma, el parámetro no resulta significativo al introducirlo en el nivel 2 (escuelas), si bien hemos comprobado que existe una correlación significativa entre esta variable y el rendimiento en matemáticas ( $r = -0,17$ ,  $p=0,000$ ), que indicaría que al aumento del número de alumnos por profesor, le corresponde una disminución del rendimiento en esta materia. Se trataría de una correlación espúrea, producida por la acción causal de una o más variables sobre el rendimiento y el nº de alumnos por aula simultáneamente, que resulta no significativa al introducir la variable en el modelo.

Respecto del *contexto educativo del centro*, las variables que han resultado significativas en el modelo no hacen sino confirmar estudios anteriores (CIDE, 2005). En primer lugar, la *existencia de cursos de refuerzo de matemáticas para aquellos alumnos que lo precisen*, supone un importante beneficio para los alumnos, ya que hace que su rendimiento aumente hasta 4. Tal y como lo ha elaborado PISA, los valores que toma la variable son 0, cuando no se ofrece ninguno; 1, cuando se ofrecen cursos de refuerzo o profundización; y 2, cuando se ofrecen de los dos tipos. Parece que el rendimiento de los alumnos aumenta en la medida en que el centro pone a su disposición los cursos que precisa.

Otro aspecto que la evidencia empírica resalta como criterio de eficacia, por su relación con el rendimiento de los alumnos, es la *frecuencia de las evaluaciones* (Stallings, 1985; García Duran, 1991). De los trabajos anteriores se deducía que una evaluación frecuente favorece un alto rendimiento, sin embargo, encontramos en nuestro modelo que, a medida que aumenta en número de evaluaciones, el rendimiento de los alumnos en matemáticas disminuye, concretamente 3 puntos por cada punto de aumento. La escala, según la construye PISA, tiene tres categorías: 0 = menos de 20 veces al año; 1 = entre 20 y 39 veces; 2 = más de 40 veces al año. Parece que, en este sentido, nuestro modelo no concuerda con investigaciones pasadas, y que cuando las evaluaciones se hacen demasiado frecuentes (más de 20 al año), el rendimiento de los alumnos desciende.

Respecto del agrupamiento de los alumnos, nuestro trabajo confirma otros como el de Slavin (1996) que llegaba a la conclusión de que la *agrupación* homogénea de

alumnos («asignación a clases de iguales aptitudes») es generalmente ineficaz. En nuestro modelo, el parámetro asociado a esta variables tiene un valor de -1,8, lo que indica que, tal y como se ha medido la variable en PISA (0 = Escuelas en las que no existe este agrupamiento; 1 = Escuelas en las que se realiza agrupamiento por habilidad en alguna clase; 2 = Escuelas en las que se agrupa por habilidad en todas las clases), a medida que se da en el centro la agrupación por habilidades, disminuye el rendimiento esperado de los alumnos en matemáticas, siendo mayor la disminución en aquellos centros en los que la agrupación por habilidad se produce en el 100% de las clases.

Por último, dentro del nivel de escuela, se confirma que la *autonomía* de ésta, en lo que a toma de decisiones se refiere, influye sobre el rendimiento de los alumnos en matemáticas. Por cada punto que aumenta la percepción que tiene el director sobre la autonomía de la escuela, aumenta el rendimiento en 1,6054 puntos. Sin embargo, es preciso tener en cuenta que los beneficios de una escuela *autónoma* no conducen de por sí al mejoramiento de la calidad. Un reciente estudio sobre la gestión basada en la escuela, concluía que «no hay prácticamente datos seguros basados en investigaciones sobre los efectos directos o indirectos que ejerce en el alumno la administración basada en la escuela. [...] Según los pocos elementos existentes basados en la investigación, los efectos en el estudiante pueden ser positivos o negativos» (Caldwell, 1998).

Por último, en el nivel *país*, hemos encontrado resultados que concuerdan, en parte, con las investigaciones que se han consultado al elaborar el marco referencial de este trabajo. El indicador que ha resultado significativo es el porcentaje de PIB que cada país invierte en educación. Esta variable, como ya se ha comentado, ha sido centrada, para facilitar su interpretación, con respecto de la media, que es del 5,1% ( $S=1$ ) (lo que pondría de manifiesto que aún no se ha conseguido el compromiso de invertir un mínimo del 6% del PIB, tal y como proponía la UNESCO, 1996). La interpretación que se hace del parámetro es que por cada punto que se sobrepasa del 5,1% del PIB invertido en educación en los países, aumenta el rendimiento medio esperado en matemáticas casi 12 puntos.

Sin embargo, otras variables que se relacionaban con la calidad educativa en otros trabajos, como el PIB del país o la inversión por niveles (Gupta, Verhoeven y Tiongsong, 1999; y McMahon, 1999) han resultado no significativas en nuestro modelo.

## Referencias bibliográficas

- BROKOVER, W.B. ET AL. (1979). *School systems and students achievement: schools make a difference*. New York: Praeger.
- CALDWELL, B. J. (1998). *Self-managing schools and improved learning outcomes*. Canberra: Department of Employment, Education, Training and Youth Affairs.
- CASTEJÓN, J.L. (1996). *Determinantes del rendimiento académico de los estudiantes y de los centros educativos: modelos y factores*. Alicante: Editorial Club Universitario.
- CIDE (1995). *El sistema educativo español*. Madrid: MEC
- (2005). *Treinta años de Investigación Educativa*. Madrid: CIDE
- COLEMAN, J.S. ET AL (1966). *Equality of educational opportunity* (2 vols.). Washington, DC: Government Printing Office.
- CREEMERS, B. (1994). *The effective classroom*. London: Cassell
- DRAPER, D. (1995). Inference and hierarchical modeling in the Social Sciences. *Journal of Educational and Behavioral Statistics*, 20, 115-147.
- EDMONDS, R.R. (1979). Effective schools for the urban poor. *Educational Leadership*, 40 (3), 4-11.
- FUENTES, A. (1986). *Procesos funcionales y eficacia de la escuela. Un modelo causal*. Madrid: Universidad Complutense.
- FULLER, B. Y CLARKE, P. (1994). Raising School Effects While Ignoring Culture? Local conditions and the influence of classrooms, tools, rules and pedagogy. *Review of Educational Research*, 1(64), 119-57.
- GARCÍA, M. (1991). *Investigación evaluativa sobre las variables pedagógicas que discriminan entre los centros de EGB de alto y bajo rendimiento en el medio rural de la provincia de Cádiz*. Tesis doctoral no publicada, UNED, Madrid, España.
- GAUTHIER, C. Y DEMBÉLÉ, M. (2004). *Qualité de l'enseignement et qualité de l'éducation*. *Revue des résultats de recherche*. Background paper for EFA Global Monitoring Report 2005.
- GAVIRIA, J.L. Y CASTRO, M. (2004). *Modelos Jerárquicos Lineales*. Madrid: La Muralla
- GIMENO, J. (1976). *Autoconcepto, sociabilidad y rendimiento escolar*. Madrid: INCIE.
- GLASMAN, N. & BINIAMINOV, I. (1981). Input-output analysis of school. *Review of Educational Research*, 51 (4), 509-539.
- GONZÁLEZ, M.C. Y TOURÓN, J. (1992). *Autoconcepto y rendimiento escolar*. Pamplona: EUNSA.

- GUPTA, S., VERHOEVEN, M. Y TIONGSON, E. (1999). *Does Higher Government Spending Buy Better Results in Education and Health Care?* Washington: International Monetary Fund. (Working Paper No 99/21, February).
- HANUSHEK, E. A. (1992). *Dropping out of school: further evidence on the role of school quality in developing countries*. Rochester: University
- HANUSHEK, E.A. Y KIMKO, D.D. (2000). Schooling, Labor-Force Quality, and the Growth of Nations. *American Economic Review*, 5(90), 1184-1208.
- LEE, J. Y BARRO, R.J. (2001). Schooling Quality in a Cross-Section of Countries. *Economica*, 272 (38), 465-488.
- LEZZOTTE, L. W., EDMONDS, R. & RATNER, G. (1974). *A final report: Remedy for school failure to equitably deliver basic schools skills*. East Lansing: Michigan State University.
- MARTÍN RODRÍGUEZ, E. (1985). *Variables de influjo inmediato en el rendimiento escolar (estudio de la incidencia de los factores socio-económicos y culturales en el rendimiento de los alumnos)*. Tesis doctoral inédita, UNED, Madrid, España.
- MCMAHON, W. (1999). *Education and Development: Measuring the Social Benefits*. Oxford: Oxford University Press.
- MUÑOZ REPISO, M. (1997). La mejora de la eficacia escolar: un estudio de casos.
- MUÑOZ REPISO, M. ET AL (1995). *Calidad de la educación y eficacia de la escuela. Estudio sobre la gestión de los recursos educativos*. Madrid: CIDE.
- MURILLO, F. J. (2000). *La Investigación sobre eficacia escolar en España*. Madrid: CIDE.
- MURPHY, J., HALLINGER, PH. Y MESA, R. (1985). School effectiveness: checking progress and assumptions and developing a role for state and federal government. *Teachers College Record*, 86 (4), 615-641.
- PÉREZ SERRANO, G. (1981). *Origen social y rendimiento social*. Madrid: CIS
- REYNOLDS, D. (1985). *Studying School Effectiveness*. London: Falmer Press.
- ROA, J.M. (2006). Rendimiento escolar y «situación diglósica» en una muestra de escolares de educación primaria en Ceuta. *Revista Electrónica de Investigación Educativa*, 8 (1).
- RODRÍGUEZ ESPINAR, S. (1979). *Factores predictivos de rendimiento escolar*. Tesis doctoral inédita. Barcelona: Universidad de Barcelona, Barcelona, España.
- RODRÍGUEZ GÓMEZ, G. (1991). *Investigación evaluativa en torno a los factores de eficacia escolar de los centros públicos de EGB*. Tesis doctoral inédita, UNED, Madrid, España.
- RUTTER, M. ET AL (1979). *Fifteen thousand hours. Secondary school and their effects on children*. London: Open Books.

- SCHEERENS, J. (1992). *Effective Schooling Research: Theory and Practice*. New York: Cassell.
- STALLINGS, J. (1985). Effective elementary classroom practices. En M.J KYLE (Ed.), *Reaching for excellence. An effective schools sourcebook*. Washington, DC: United States Government Printing Office
- TEDDLIE, C. Y STRINGFIELD, S. (1993). *Schools make a difference: lessons learned from a ten-year study of school effects*. New York: Teachers College Press.
- UNESCO (1996). *International Standard Classification of Education*. Paris: UNESCO
- UNESCO (2005). *Informe de Seguimiento de la Educación para Todos en el Mundo*. Paris: Autor
- WALBERG, H. (1991). Improving school science in advanced and developing countries *Review of Educational Research*, 1(60), 25-69.
- WU, M. (2002). *Manual de análisis de datos de PISA 2003: usuarios de SPSS*.

## Fuentes electrónicas

- GLEWWE, P., KREMER, M. Y MOULIN, S. (2000). *Textbooks and Test Scores: evidence from the Prospective Evaluation in Kenya Working Paper*. Cambridge, MA: Harvard University. Consultado el 10 de noviembre, de <http://posteconomics.harvard.edu/faculty/kremer/webpapers/>
- HORSLEY, M. (2004). *An Expert Teacher 's Use of Textbooks in the Classroom*. University of Sydney, Consultado de [http://alex.edfac.usydedu.au/Year1/cases/Case%2014/Expert\\_teacher's\\_use\\_of\\_te.html](http://alex.edfac.usydedu.au/Year1/cases/Case%2014/Expert_teacher's_use_of_te.html)
- OECD (2003). *The PISA International Database*. Autor: OECD, de <http://pisa2003.acer.edu.au/index.php>
- OECD (2004). *Education at glance*. París: ECD. Consultado de [www.oecd.org/edu/eag2004](http://www.oecd.org/edu/eag2004)

**Dirección de contacto:** Covadonga Ruiz de Miguel. Universidad Complutense de Madrid. Facultad de Educación. Departamento de Métodos de Investigación y Diagnóstico en Educación. Edificio La Almudena, c/ Rector Royo Villanova s/n. Ciudad Universitaria. 28040, Madrid, España. E-mail: covaruiz@edu.ucm.es

# Diseño de propuestas docentes con TIC para la enseñanza de la autorregulación en la Educación Superior

## Instructional design with ICT for self-regulated learning in Higher Education

Teresa Mauri Majós, Rosa Colomina Álvarez e Inés de Gispert

*Universidad de Barcelona. Departamento de Psicología Evolutiva y de la Educación*

### Resumen

Las exigencias de innovación de la docencia universitaria derivadas de los planteamientos del Espacio Europeo de Educación Superior sitúan la enseñanza y el aprendizaje de las competencias en el centro mismo de la tarea docente y, en especial, la competencia de autorregulación del aprendizaje. Así mismo, plantean que, en la mayoría de las ocasiones, elaborar una respuesta eficaz a los nuevos retos comporta que la tarea de innovación deba ser abordada colectivamente por un equipo de profesores. Este trabajo presenta un estudio sobre una propuesta de innovación para la enseñanza de las competencias de autorregulación del aprendizaje individual y de grupo, diseñada y desarrollada en colaboración por un equipo docente. Los datos analizados corresponden a seis grupos-clase (un total de seis profesores y 283 alumnos) de una asignatura de Psicología de la Educación que se imparte en la Facultad de Formación de Profesorado, en la cual se integra el proceso de enseñanza y aprendizaje de la autorregulación. Tras una breve reflexión sobre la caracterización de las competencias, en primer lugar se presentan las opciones adoptadas en el diseño de la propuesta que se fundamenta en una perspectiva socioconstructivista de los procesos instruccionales, y se detallan las dimensiones de autorregulación que son objeto específico de este estudio. En segundo lugar, se presentan los resultados relativos al aprendizaje de la autorregulación de los alumnos, y también relativos a la experiencia de colaboración desarrollada por los profesores. Finalmente, unas breves conclusiones enfatizan los puntos importantes que deben considerarse en la enseñanza de las

competencias de autorregulación para el aprendizaje y la relevancia del desarrollo de ese proceso para el profesorado.

*Palabras clave:* diseño instruccional, competencias, autorregulación del aprendizaje, ayuda educativa, perspectiva socioconstructivista, usos de las TIC en educación, aprendizaje colaborativo, evaluación de la calidad de la innovación, enseñanza universitaria.

### **Abstract**

The new European Space for Higher Education places new challenges for innovation in university education and brings the teaching of competences to the core of the teaching task, especially the competence of self-regulated learning. These new circumstances also highlight the need for an efficient educational response to these challenges, which often implies the collaboration of a team of faculty members. This article presents a study on a teaching innovation experience and it focuses on developing competences of self-regulated learning, both individual and team learning. It was designed and implemented by a team of teacher educators. The data belong to a set of six classes (six teachers and 283 students) of a course of Educational Psychology at a Teacher Training School, which combines the process of self-regulated teaching and learning. Following a reflection on characterization of competences, we firstly present the criteria and the design decisions taken for our proposal, based on a socio-constructivist perspective of instructional processes. We also present the dimensions of self-regulation, which are the focus of the study. Then we present the results concerning the students' learning of self-regulation and also, those related to the teachers' experience in the collaborative design of teaching. Finally, a number of brief conclusions are drawn to emphasize the most important points which need to be taken into account when teaching competences for self-regulated learning, as well as the relevance of this process for the teaching community.

*Key Words:* instructional design, higher education, competences, self-regulated learning, educational guidance, socio-constructivist perspective, uses of ICT in education, collaborative learning, assessment of teaching innovation quality, higher education.

El Espacio Europeo de Educación Superior ha convertido las competencias en un referente importante para la elaboración de los planes de estudio y las propuestas docentes de las diferentes asignaturas de las titulaciones universitarias. Este trabajo presenta los resultados de una investigación sobre una experiencia de innovación

docente<sup>1</sup> en relación con la enseñanza y aprendizaje de las competencias de autorregulación en la educación superior.

El término *competencia* parece tener un carácter polisémico y es difícil encontrar referencias en un sentido unívoco en los documentos-marco dirigidos al profesorado universitario (Esteve Zarazaga et al., 2006). Sin embargo, para nuestros propósitos, destacaremos que el concepto de competencia pone el acento sobre los contenidos o conocimientos –de diferente naturaleza– imprescindibles para llevar a cabo actividades relevantes y significativas en contextos variados y funcionales. Desde esta perspectiva, toda competencia integra la combinación de saberes de muy distinto tipo y la capacidad de movilizarlos de forma consciente en los contextos o situaciones cultural, académica y profesionalmente relevantes en los que los estudiantes deben aprender a actuar. Por lo tanto, disponer de diferentes tipos de conocimientos es una condición necesaria de la caracterización de la competencia, aunque no suficiente: exige contar también con la posibilidad de utilizar esos conocimientos para actuar de forma consciente en distintos contextos (Perrenoud, 2004). El concepto *consciente* es usado en este artículo desde una perspectiva vygotskyana que señala que los procesos psicológicos superiores guían la conducta humana a través de la autorregulación permitiendo una adaptación activa, transformadora de la realidad y de las propias competencias (Vygotsky, 1925). Así, la autorregulación tiene un papel central en el aprendizaje de otras competencias y, en consecuencia, se considera que debe estar presente en la elaboración de propuestas docentes innovadoras que incluyan ayudas específicas dirigidas a orientar, sostener y guiar su aprendizaje.

La presentación del trabajo se inicia, en primer lugar, con una breve referencia a los criterios y opciones psicopedagógicos que han vertebrado el diseño y desarrollo de la propuesta docente para la enseñanza de la autorregulación. En segundo lugar, se presentan los resultados relativos al aprendizaje de la autorregulación por los alumnos, y a la valoración de la experiencia por parte de los profesores. Finalmente, cerraremos nuestra aportación con unas breves conclusiones.

---

<sup>1)</sup> La realización de este trabajo forma parte de una línea de investigación sobre docencia desarrollada por el *Grupo de innovación docente en Psicología de la Educación* (GIDPE) de la Universidad de Barcelona. El GIDPE está dirigido por el Dr. César Coll y formado por Rosa Colomina, Anna Engel, Inés de Gispert, Teresa Mauri, Rosa Mayordomo, Javier Onrubia, Ana Remesal, M<sup>a</sup> José Rochera y Teresa Segué. En el desarrollo de este proyecto, confluyen dos de las líneas prioritarias de trabajo que han configurado la dedicación del grupo en los últimos años: la investigación sobre los usos psicopedagógicos de las TIC en los procesos de enseñanza y aprendizaje y el diseño y desarrollo de procesos de innovación docente en la enseñanza universitaria de asignaturas de Psicología de la Educación que se imparten en la Universidad de Barcelona.

## El aprendizaje y la enseñanza de la autorregulación: una experiencia de innovación docente basada en principios constructivistas

El diseño de la innovación de la enseñanza de la autorregulación se aborda desde unos criterios psicopedagógicos que responden, por una parte, a una perspectiva constructivista de orientación sociocultural de los procesos de enseñanza y aprendizaje (Coll 2001; 2004) y, por otra, a la reflexión continuada del equipo de profesores sobre la práctica docente y los planteamientos sobre la integración de la enseñanza de la autorregulación en la asignatura<sup>2</sup>. Desde estas coordenadas, se orientó el diseño de la propuesta de ayudar a los alumnos a mejorar sus conocimientos y habilidades en relación con los procesos de autorregulación en general, y de los procesos de planificación del aprendizaje individual y de grupo en particular. El aprendizaje de la autorregulación se focalizó en la progresiva asunción de responsabilidad por parte del estudiante en la planificación, el control de su proceso de aprendizaje y la evaluación de su efectividad para la consecución de los objetivos propuestos (Arvaja et al. 2007; Perry et al, 2002; Pintrich y Zusho, 2002; Pozo y Monereo, 2002; Zeidner et al. 2000; Schunk y Zimmerman, 2003).

La propuesta docente se centró en enseñar a los alumnos a distribuir y gestionar el tiempo de aprendizaje individual y de grupo, con el propósito de que aprendieran a tomar en consideración las características, las exigencias y las condiciones de realización y de utilización de los recursos disponibles para el desarrollo de las tareas académicas. Las dimensiones que se consideraron en relación con la autorregulación se presentan en la Tabla 1.

**TABLA I.** Dimensiones de análisis de la autorregulación en el proceso de enseñanza y aprendizaje.

1. Representarse la tarea de aprendizaje en su conjunto
2. Valorar las habilidades y los conocimientos previos de quienes participan en el desarrollo de la tarea como individuos y como grupo: puntos fuertes y puntos débiles
3. Identificar las condiciones de desarrollo de la tarea: plazos, materiales, formas de organización del trabajo, acceso a la información, estrategias y técnicas de realización más adecuadas
4. Planificar la actividad de aprendizaje a desarrollar: distribución del tiempo, individualmente y en grupo, tomando en consideración los diferentes aspectos y las condiciones del proceso de enseñanza y aprendizaje
5. Usar la planificación elaborada a lo largo del proceso de enseñanza y aprendizaje para guiar la actividad individual y conjunta y ajustarla, si cabe, a las nuevas condiciones y necesidades que aparecen en el proceso
6. Valorar el proceso y concretar la estrategia a seguir en las futuras realizaciones de tareas semejantes a la desarrollada

<sup>(2)</sup> Las ideas y planteamientos de este trabajo son deudoras del trabajo desarrollado por el GIDPE en el proyecto sobre *El aprendizaje de la autorregulación individual y de grupo. Una propuesta centrada en el trabajo del alumno y el uso de las TIC en el ámbito de la Psicología de la Educación*, subvencionado por el DURSI (SGR-2005MQD00149), así como del proyecto de investigación sobre docencia *Instrumentos para la evaluación de procesos de innovación docente universitaria desde una perspectiva constructivista*, subvencionado por el REDICE-04.

Los datos sobre la situación de enseñanza y aprendizaje de la autorregulación se recogen en la asignatura de Psicología de la Educación y Atención a la Diversidad (anual, de carácter troncal y 12 créditos ECTS) que se imparte en la facultad de formación del profesorado de la Universidad de Barcelona. La intervención específica del profesorado para el aprendizaje de la autorregulación se desarrolló durante el trabajo de los tres últimos bloques de los seis que componen la asignatura, en los que los estudiantes trabajan contenidos sobre diferentes factores que intervienen en el aprendizaje escolar mediante la metodología de análisis de casos y con el uso de TIC (en concreto, el uso de la plataforma Moodle). Han participado en esta experiencia seis grupos-clase con un total de 283 estudiantes y seis profesores.

Para el análisis de la experiencia de innovación en relación con los estudiantes se recogieron las informaciones siguientes: datos sobre su concepción inicial de la autorregulación y el papel atribuido a la misma en el aprendizaje, y datos de la valoración de su uso en la regulación del aprendizaje individual y de grupo; datos sobre el uso real y las modificaciones de la planificación por parte de los estudiantes durante el proceso de enseñanza y aprendizaje; datos sobre su concepción final de la autorregulación y el papel atribuido a la misma en el aprendizaje, y datos de la valoración de su uso en la regulación del aprendizaje individual y de grupo. Para ello se elaboraron, por una parte, un cuestionario cuyo reparto se llevó a cabo al inicio y al final del proceso y, por otra, una pauta de seguimiento de la elaboración y uso de la planificación que se utilizó al finalizar las dos primeras experiencias de uso de la planificación por parte de los estudiantes (al final de los bloques IV y V). También se recogieron para su análisis las planificaciones elaboradas por los alumnos.

A continuación, se presentan las principales opciones que fundamentan el diseño y desarrollo de la propuesta docente: las relaciones entre las actividades de enseñanza y aprendizaje y la autorregulación; la enseñanza y aprendizaje de la autorregulación como sistema vertebrado en torno a bloques temáticos, y el ajuste de la ayuda educativa al aprendizaje de la autorregulación. Cabe señalar que en los tres tipos de opciones se incluyen criterios que tienen consecuencias para la ayuda educativa prestada por el profesor al alumno si bien ésta se concreta, en cada caso, mediante diferentes recursos y en diferentes momentos del proceso de enseñanza y aprendizaje.

## La autorregulación y su incorporación en las actividades de enseñanza y aprendizaje

La autorregulación se aborda en este estudio como un proceso inherente al proceso mismo de enseñanza y aprendizaje, que necesita de espacios de reflexión y de actuaciones o intervenciones educativas específicas del profesor para que el alumno progrese en el dominio de estas competencias. De acuerdo con este criterio, las actividades de enseñanza y aprendizaje se conciben, al mismo tiempo, como actividades de enseñanza y aprendizaje de la autorregulación. Para ilustrar este hecho, la Tabla II presenta tipos diferentes de actividades de enseñanza y aprendizaje de contenidos de psicología de la educación que incorporan la metodología de análisis de casos, mostrando, además, la doble función que cumplen de ser útiles al profesor para enseñar a autorregular el aprendizaje individual y de grupo, y al alumno para aprender a autorregular dicho proceso.

**TABLA II.** Actividades de enseñanza y aprendizaje y su función para la enseñanza y el aprendizaje de la autorregulación

Actividades de enseñanza y aprendizaje	Función para la enseñanza de la autorregulación	Función para el aprendizaje de la autorregulación
Actividades de presentación del caso y de presentación de las actividades de enseñanza y aprendizaje relacionadas	Presentar los objetivos, el caso y la tarea. Se trata de dar a conocer las exigencias de las diferentes actividades para analizar las condiciones de realización e identificar los elementos y materiales de apoyo	Representación inicial de las exigencias de análisis del caso y del papel de cada una de las actividades en el análisis y resolución del mismo  Identificar las condiciones de realización de las actividades: plazos, materiales, formas de organización del trabajo, acceso a la información, estrategias y técnicas más adecuadas para realizar la tarea.
Actividades para hacer emerger los conocimientos y las habilidades de los alumnos para abordar el caso	Recoger información de los conocimientos y las habilidades iniciales de los alumnos, incluyendo los conocimientos y habilidades relativos a la autorregulación  Realizar una evaluación inicial regulativa pro-activa de los distintos saberes del alumnado	Valorar las habilidades y los conocimientos propios necesarios para la actividad  Identificar la necesidad de profundizar en la comprensión del contenido y de planificar el trabajo
Actividades de elaboración significativa de nuevos conocimientos  Actividades de uso funcional y contextualizado de los conocimientos adquiridos en contextos nuevos simulados	Recoger información continua y sistemáticamente del dominio progresivo por el alumno de los contenidos y de las habilidades de análisis del caso  Recoger información de las necesidades educativas del alumno para ajustar la ayuda al proceso de aprendizaje y de autorregulación del aprendizaje	Seguir la planificación elaborada a lo largo del proceso de enseñanza y aprendizaje y ajustarla, si cabe, a las nuevas condiciones y necesidades que aparecen en el proceso  Decidir demandas de ayuda al profesor  Elaborar diferentes productos vinculados al análisis del caso

Actividades de enseñanza y aprendizaje	Función para la enseñanza de la autorregulación	Función para el aprendizaje de la autorregulación
Actividades de evaluación final y de autoevaluación individual y de grupo	Realizar una actividad reguladora final de apoyo para hacer emerger los procesos metacognitivos implicados en la autorregulación	Retomar el análisis del caso o el problema planteado a través del uso contextualizado de los conocimientos adquiridos  Reflexionar y tomar conciencia del proceso realizado y concretar la estrategia a seguir en la realización futura de tareas semejantes

Estas actividades potencian el desarrollo de formas de actuación consciente y autorregulada para favorecer la autonomía y la responsabilidad progresiva de los alumnos en su aprendizaje.

### La enseñanza y aprendizaje de la autorregulación como sistema vertebrado en torno a bloques temáticos

Para la comprensión de la enseñanza de la autorregulación en esta experiencia es esencial destacar que el trabajo de la asignatura se planifica en unidades de contenido amplias que se denominan bloques temáticos, criterio que responde a la finalidad de favorecer un aprendizaje funcional, situado y complejo por parte de los estudiantes. Dichos bloques responden a núcleos de contenido cuyo aprendizaje contribuye de modo significativo al desarrollo de las competencias específicas de la titulación y de otras competencias más transversales, como las de autorregulación.

La secuencia de bloques se establece siguiendo criterios propios del aprendizaje significativo: a partir de un bloque inicial ubicado a principio de curso, que contiene un núcleo de conceptos básicos, muy generales y de fácil comprensión para el alumno, se progresa para abordar otros bloques que permiten ampliar los contenidos, profundizar en ellos y reelaborar las relaciones iniciales establecidas de modo que resulten progresivamente más complejas y significativas (Mauri et al., 2005; Onrubia, 2005; Coll et al., 2006).

Cada bloque temático incluye actividades de aprendizaje y enseñanza y de evaluación dirigidas a lograr la comprensión del conocimiento y su aplicación en contextos reales simulados de análisis de casos, y la comprensión y uso progresivo de la autorregulación (Jonassen y Hernández-Serrano, 2002). Así, *las actividades de aprendizaje* cumplen un doble objetivo: hacer avanzar al alumno en la construcción de conocimientos muy variados (conceptos, procedimientos y actitudes, valores y normas) en

el marco del análisis de un caso; y crear situaciones dirigidas a aprender a usar estos conocimientos y a gestionar su proceso de aprendizaje. Para ello, se han diseñado tres tipos de actividades secuenciadas: la primera actividad sirve para descubrir los conocimientos previos de los alumnos sobre el tema; la segunda actividad se dirige a ayudar al alumno a establecer relaciones significativas entre esos conocimientos previos y la nueva información que se le presenta. La tercera actividad tiene como finalidad contribuir a dotar de funcionalidad los aprendizajes adquiridos, dando al alumno la oportunidad de usarlos en un contexto de práctica concreto. Por su parte, *las actividades de evaluación* se han elaborado para regular el aprendizaje y valorar el aprendizaje alcanzado por los alumnos. En conjunto, las diferentes actividades permiten tanto el aprendizaje de los conocimientos nucleares y relevantes del bloque como la recogida de información sobre el proceso de aprendizaje y de autorregulación de los alumnos.

Cada *bloque temático* sigue una estructura estable con distintas fases, cada una de las cuales se desarrolla mediante un conjunto también estable de elementos y recursos que se sintetizan a continuación:

■ *Presentación del bloque temático a los alumnos*

La guía docente del bloque es el recurso que contribuye a que profesor y alumnos elaboren una definición inicial compartida de los objetivos y las actividades del bloque que les permita orientar el aprendizaje y elaborar el plan de trabajo adecuado para llevarlo a cabo. Para apoyar a los alumnos en ese propósito, junto con la guía docente, se presenta una *agenda de trabajo recomendada*.

■ *Desarrollo del bloque*

El desarrollo del bloque integra habitualmente sesiones de aportación de información y otras de desarrollo de actividades de enseñanza y aprendizaje. Son recursos-tipo de esta fase: los *esquemas del contenido del bloque*, las *pautas de lectura* o documentos de ayuda a la realización de una lectura comprensiva de los textos del curso, y las orientaciones para las actividades de aprendizaje. El diseño de estas *orientaciones para las actividades* de aprendizaje responde no sólo a la necesidad de concretar la demanda de la tarea, sino también de contribuir a guiar adecuadamente el esfuerzo del alumno y servir de regulación para el desarrollo de la actividad. Por ello, cada una de las actividades se ha diseñado integrando los siguientes elementos: los objetivos de la actividad propiamente dicha, el producto final que debe ser logrado como resultado de la actividad, las recomendaciones para ayudar al

alumno o al grupo de alumnos a planificar y desarrollar la actividad, y los criterios de evaluación.

#### ■ *Final del bloque*

El final del bloque incluye la síntesis de los contenidos y las tareas de *evaluación final* del bloque. Durante esa fase, el profesor presenta una síntesis de los conocimientos adquiridos a lo largo del bloque. La finalidad es identificar los más relevantes y mostrar, de nuevo de forma explícita, las relaciones existentes entre ellos. Mediante este modo de proceder, se consigue que dichos saberes puedan ser identificados por todos los implicados, profesores y alumnos, como «el conocimiento que se ha construido y que se comparte», y que los alumnos tengan una nueva oportunidad para identificar y resolver dudas al respecto.

La secuencia de actividades de enseñanza y aprendizaje que acabamos de describir proporciona a los estudiantes oportunidades muy variadas para abordar autónomamente situaciones de aprendizaje complejas: planificar su actuación; decidir qué conocimientos se han de utilizar y cómo se han de utilizar; comparar el planteamiento inicial del aprendizaje y el abordaje inicial de las actividades basado en sus conocimientos previos con la resolución final de las tareas o problemas; reflexionar sobre el curso de la acción seguido y reorientar el proceso de aprendizaje.

Asimismo, todos estos aspectos y el orden que sigue la secuencia en todos los bloques temáticos permiten al profesor obtener múltiples informaciones del proceso de aprendizaje de los alumnos y del modo como ejercen la autorregulación, por lo que es posible el seguimiento y el ajuste continuado de la ayuda para mejorar las competencias de autorregulación. En concreto, la intervención específica del profesorado se desarrolla a lo largo de los tres últimos bloques de la asignatura y se diseñó para proceder desde un alto grado de intervención sobre la planificación en el bloque el que se inició la experiencia (bloque IV), a un menor grado de intervención en el siguiente bloque temático (bloque V) hasta retirarla en el bloque temático final del curso (bloque VI). A saber, el profesorado varió la forma en la que orientó, sostuvo y ayudó al alumno y al grupo de alumnos a planificar y a usar la planificación en la regulación y control del aprendizaje a lo largo de los dos bloques temáticos iniciales y se abstuvo de hacer lo propio en el tercero.

A lo largo de los dos primeros bloques temáticos, las diferentes dimensiones de la autorregulación fueron objeto de presentación, estudio y debate gracias a la actuación del profesorado basada en las medidas siguientes: mediante el reparto y devolución

de los resultados del cuestionario inicial a los alumnos, que facilitó la reflexión sobre la concepción y el proceso de autorregulación y su papel en el aprendizaje; aportando modelos de planificación diferentes y valorando sus cualidades; potenciando el uso de la planificación en la regulación del trabajo individual y de grupo, y favoreciendo la integración de cambios para ajustarla a las nuevas demandas y condiciones de realización de la tarea, y explicitando el proceso de regulación del trabajo individual y de grupo basado en el uso de la planificación. Durante el tercer bloque, los alumnos no recibieron ninguna demanda explícita de planificación. Dado que el propósito era identificar los avances en el uso autónomo de la planificación para la autorregulación, en el bloque VI únicamente se recogió la planificación efectuada por aquellos alumnos que la habían llevado a cabo espontáneamente.

### **El papel del profesor: el ajuste de la ayuda educativa al aprendizaje de la autorregulación**

Según lo que se ha comentado en apartados anteriores, la autorregulación no es una competencia que el alumno posea de entrada o que le sea fácil de adquirir si se relega a la espontaneidad de las reflexiones que el estudiante pueda realizar por sí mismo. En este sentido, la estructura del bloque, las demandas que genera y el conjunto de instrumentos y recursos a disposición del alumnado y del profesor permiten crear un contexto óptimo para el ajuste de la ayuda educativa facilitando que el profesor pueda identificar la actuación que el alumno realiza para autorregular el aprendizaje y proporcionar la orientación y apoyo necesarios para mejorar dicha actividad reguladora.

La relación entre las dimensiones de la autorregulación y los recursos de ajuste de la ayuda educativa para mejorar estas competencias se muestra en la Tabla III.

**TABLA III.** Relación entre las dimensiones de la autorregulación y los recursos de ajuste de la ayuda educativa

Dimensiones de la autorregulación	Recursos de ayuda educativa al aprendizaje de la autorregulación
a. Representación de la actividad de enseñanza y aprendizaje	<b>1- Relativo al bloque</b> 1.1. Guía didáctica del bloque 1.2. Calendario de trabajo del bloque
c. Identificación de las condiciones de la tarea	<b>2- Relativo a la actividad o tarea</b> 2.1. Presentación de la/s tarea/s o actividad/es 2.2. Pautas de lectura 2.3. Material complementario sobre la actividad

Dimensiones de la autorregulación	Recursos de ayuda educativa al aprendizaje de la autorregulación
b. Valorar e identificar las habilidades, los conocimientos y recursos disponibles	Pauta para identificar las habilidades y los conocimientos de los que dispone el alumno individualmente y el grupo de trabajo para desarrollar la planificación
d. Planificación propiamente dicha	Actividad de planificación: Pauta-guía de elaboración de la planificación. Incluye ejemplos y modelos de planificación.
e. Seguimiento de la planificación	Actividad de uso y revisión de la planificación.
f. Valoración	<ol style="list-style-type: none"> <li>1. Reflexión conjunta de profesor y alumnos sobre las respuestas dadas a un cuestionario sobre las dimensiones de la autorregulación (ver Tabla I) y su uso en el proceso de enseñanza y aprendizaje.</li> <li>2. Puesta en común del proceso de enseñanza y aprendizaje.</li> <li>3. Autoevaluación del aprendizaje individual y de grupo.</li> </ol>

Se han diseñado, asimismo, instrumentos de seguimiento del aprendizaje del alumno, como el espacio dedicado a la *tutoría online*, que pretende conseguir que los malentendidos o incomprensiones sobre el contenido y sobre los aspectos organizativos se aborden desde el mismo momento en que el alumno lo necesite para evitar que repercutan negativamente en su implicación activa y continuada en el aprendizaje. Además, se han elaborado instrumentos de *autoevaluación del aprendizaje individual y de grupo* para que los alumnos reflexionen sobre el proceso y los resultados de su aprendizaje y decidan las mejoras que es preciso llevar a cabo en el futuro. Finalmente, el equipo docente ha diseñado un espacio para favorecer la comunicación entre alumnos en el trabajo colaborativo, que incluye diferentes herramientas para el uso exclusivo de los grupos: *el forum y el wiki*. En este sentido, como se mostrará más adelante, la incorporación de las TIC constituye un instrumento privilegiado para que el profesorado pueda seguir los procesos de trabajo en grupo que le resultaban, hasta hace poco tiempo, opacos y ajustar, en consecuencia, la ayuda educativa en el momento en el que surgen dudas o dificultades en el trabajo en grupo.

## Resultados de la experiencia: mejora de la autorregulación en el caso de los estudiantes y valoración de las consecuencias de la colaboración docente en el caso de los profesores

En este apartado se presentan, por una parte, los resultados de los alumnos en relación con la autorregulación y, por otra, la valoración del profesorado implicado en la experiencia.

## Resultados de la valoración de los estudiantes sobre distintos elementos de la autorregulación

En relación con el análisis de los datos referidos a los alumnos, se presentan, en primer lugar, los resultados de los cuestionarios sobre la valoración y las creencias sobre la autorregulación. En segundo lugar, se presentan algunos datos evolutivos identificados en el análisis comparativo de las respuestas de una pauta sobre el uso de la planificación, usada en dos momentos diferentes de la experiencia. En tercer lugar, se revisan algunos de los resultados del análisis de los documentos de planificación elaborados por los alumnos a lo largo de los tres bloques.

### Resultados del cuestionario de valoración de la experiencia de los estudiantes

El cuestionario consta de 20 preguntas sobre las distintas dimensiones de la autorregulación que se plantean desde la perspectiva de la planificación, y presentan posibilidad de respuesta graduada en seis opciones: «muy de acuerdo», «de acuerdo», «un poco de acuerdo», «un poco en desacuerdo», «en desacuerdo» y «muy en desacuerdo». Los resultados del cuestionario se describen a continuación tomando como referencia las mismas dimensiones de estudio y análisis. Para facilitar la interpretación de los datos se destacan las tendencias de las respuestas de los estudiantes (expresadas en porcentajes) que mejor permiten dotar de significado a los resultados obtenidos, aglutinando para ello, en algunos casos, más de una opción de respuesta.

- Dimensión relativa a *Representarse la tarea de aprendizaje* en su conjunto. Los estudiantes se muestran mayoritariamente «de acuerdo» o «muy de acuerdo» (81,27%) en dedicar un tiempo del trabajo personal a conocer y concretar los objetivos que guían la tarea; de modo muy relacionado con ello, los estudiantes señalan estar «de acuerdo» o «muy de acuerdo» en el interés que tiene para regular el aprendizaje el hecho de conocer las ideas del profesor sobre la tarea o actividad (68,55%), conocer la representación que tienen los compañeros del grupo sobre las actividades (68,9%), y representarse y acordar conjuntamente con los compañeros de grupo los objetivos de aprendizaje de las actividades (68,19%).
- Dimensión relativa a *Valorar las habilidades y los conocimientos previos de quienes participan en el desarrollo de la tarea como individuos y como grupo*. Los estudiantes se muestran mayoritariamente «de acuerdo» o «muy de acuerdo» (76,67%) en que ser conscientes de lo que saben y de lo que no saben

sobre el proceso de realización de la actividad incide en el aprendizaje que pueden llegar a realizar; sin embargo, es menor el número de alumnos que se muestra «de acuerdo» o «muy de acuerdo» (58,30%) en la necesidad de dedicar tiempo a relacionar los objetivos de la actividad con los propios intereses.

■ Dimensión relativa a *Identificar las condiciones de desarrollo de la tarea*.

Los estudiantes se muestran mayoritariamente «de acuerdo» o «muy de acuerdo» (71,73%) en que conocer estas condiciones facilita el aprendizaje. Sin embargo, el grado de acuerdo sobre la necesidad de calcular el tiempo, el esfuerzo y la dificultad de las actividades para facilitar el aprendizaje es menor y la dispersión de las respuestas es mayor (un 46,99% se muestra «de acuerdo» o «muy de acuerdo»; un 31,8% se muestra «un poco de acuerdo», y un 10,95% se muestra «un poco en desacuerdo»). Los estudiantes se muestran mayoritariamente «de acuerdo» o «muy de acuerdo» (69,96%) en la necesidad de dedicar tiempo a elegir la mejor estrategia de aprendizaje en cada situación.

■ Dimensión relativa a *Planificar la actividad de aprendizaje*.

Los estudiantes muestran diversidad de valoraciones en cuanto a la necesidad de elaborar una planificación por escrito (un 33,21% se muestra «un poco de acuerdo»; un 27,91% se muestra «de acuerdo»; un 20,49% se muestra «un poco en desacuerdo» o «en desacuerdo», y un 10,29% se muestra «muy de acuerdo»). Sin embargo, ante una mayor complejidad de la tarea, los estudiantes expresan una mayor necesidad de detallar por escrito la planificación (un 66,43% se muestra «un poco de acuerdo» o «de acuerdo»; un 13,07% «un poco en desacuerdo» o «en desacuerdo», y 12,01% «muy de acuerdo»). Los estudiantes atribuyen al hecho de tener experiencia previa menor importancia para elaborar una planificación detallada y por escrito (un 62,54% se muestra «un poco de acuerdo» o «de acuerdo»; un 19,43% se muestra «un poco en desacuerdo» o «en desacuerdo», y un 9,54% «muy de acuerdo»).

■ Dimensión relativa a *Usar la planificación elaborada a lo largo del proceso de enseñanza y aprendizaje para guiar la actividad individual y conjunta y ajustarla*.

Los estudiantes se muestran «de acuerdo» o «un poco de acuerdo» (69,61%) a consultar regularmente la planificación elaborada en la realización de las actividades y favorables a la introducción de los cambios necesarios en el plan de trabajo durante el proceso para conseguir un mejor ajuste a las necesidades del aprendizaje (un 61,13% se muestra «de acuerdo» o «muy de acuerdo», y un 21,90% se muestra «un poco de acuerdo»).

■ Dimensión relativa a *Valorar el proceso y concretar la estrategia que se ha de seguir en tareas semejantes.*

El acuerdo entre los estudiantes en dedicar tiempo a revisar las decisiones adoptadas durante el proceso de trabajo y a valorar su eficacia es mayoritario. Dicho acuerdo se distribuye del modo siguiente: un 55,83% «de acuerdo»; un 20,14% «un poco de acuerdo» y un 12,36% «muy de acuerdo». Sin embargo, el acuerdo en dedicar tiempo a la revisión de la planificación se muestra de modo diferente entre los estudiantes (únicamente un 63,60% está «un poco de acuerdo» o «de acuerdo»). Finalmente, dedicar tiempo a decidir cómo actuar en una situación que se desarrolla en las mismas condiciones que otra desarrollada anteriormente cuenta con un 71,37% de estudiantes que se muestran «un poco de acuerdo» o «de acuerdo» y un 12,01% «muy de acuerdo».

Un último dato hace referencia a que un porcentaje de estudiantes, que oscila entre un 7,4 % y un 9,18%, no contesta las preguntas planteadas, lo que podría estar expresando también la dificultad que supone abordar el tema de las competencias de planificación para una parte de los alumnos.

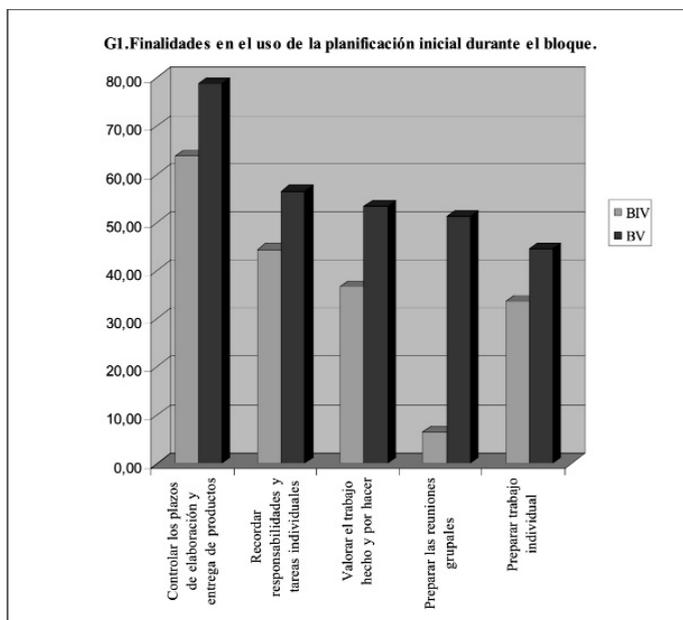
En conjunto, los estudiantes muestran una alta valoración de los aspectos relacionados con dedicar tiempo y esfuerzo a representarse la tarea y valorar las habilidades y los conocimientos previos de quienes participan en el desarrollo de la misma. Sin embargo, la elaboración y revisión por escrito de la planificación y su uso continuado a lo largo del proceso son aspectos menos valorados por parte de los estudiantes. Muy probablemente ello sea debido, en parte, a que estos aspectos de la planificación les plantean dificultades, como indicarán otros resultados más adelante.

### **Resultados del uso de la pauta de seguimiento de la planificación y su valoración**

La pauta pretende valorar el uso real que los estudiantes realizaron de la planificación a lo largo del proceso de enseñanza y aprendizaje de los bloques objeto de estudio y consta de 15 preguntas -en su mayoría de tipo cerrado y algunas de tipo abierto-. Los resultados muestran, en primer lugar, que el porcentaje de alumnos que ha usado la planificación supera el 80% y que más de un 75% de los estudiantes valora «mucho» o «bastante» la contribución de la planificación a la regulación y control del aprendizaje.

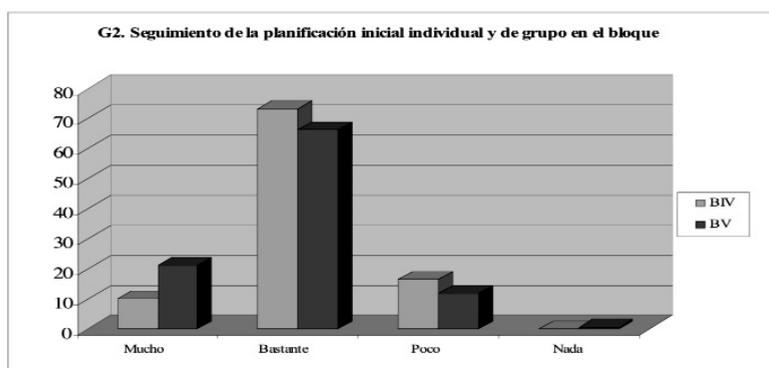
El Gráfico I detalla los distintos usos de la planificación en los dos primeros bloques de los tres que abarca la experiencia, y el incremento del porcentaje de uso de la planificación en cada uno de ellos en el segundo bloque.

GRÁFICO I. Finalidades en el uso de la planificación inicial durante el bloque



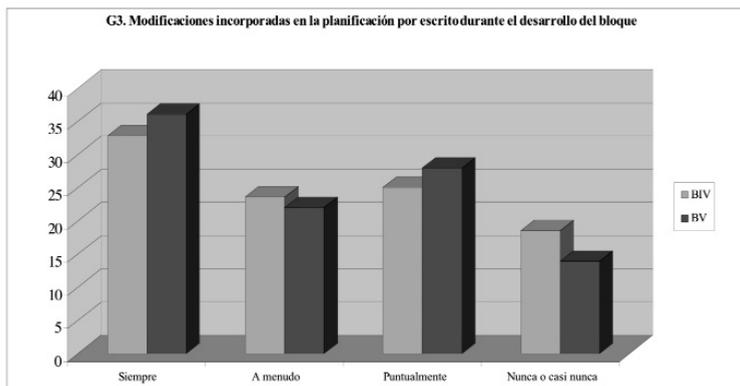
El seguimiento global de la planificación es muy similar en ambos bloques: más de un 75% de los estudiantes afirma que sigue el plan de trabajo previsto «mucho» o «bastante». El Gráfico II muestra que el porcentaje de estudiantes que escogen la opción «mucho» se incrementa en el segundo de los bloques temáticos de la experiencia.

GRÁFICO II. Seguimiento de la planificación inicial individual y de grupo durante el bloque



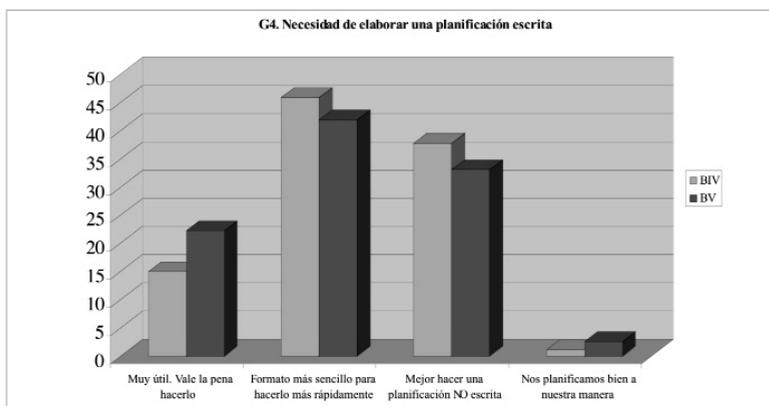
Los estudiantes que introducen modificaciones en su planificación escrita sistemáticamente aumentan en el segundo bloque y, al mismo tiempo, disminuyen los estudiantes que afirman no introducirlas nunca (véase Gráfico III).

**GRÁFICO III.** Modificaciones incorporadas en la planificación escrita durante el bloque



Como ilustra el Gráfico IV, la mayoría de estudiantes se expresa a favor de realizar la planificación por escrito de modo que requiera menos inversión de esfuerzo y tiempo que el empleado en la experiencia. No obstante, se observa un cierto incremento de los estudiantes que se muestran a favor de planificar por escrito con detalle. La tendencia de algunos estudiantes a no elaborar la planificación escrita se mantiene y, aunque decrece progresivamente en ambos bloques, al final de la experiencia es de un 30%.

**GRÁFICO IV.** Necesidad de elaborar una planificación escrita



En síntesis, los estudiantes, progresivamente a lo largo de los dos primeros bloques, han elaborado la planificación y han realizado un elevado uso de la misma para regular el aprendizaje y han valorado positivamente su contribución al mismo. Además, a medida que se muestran más expertos en el uso de la planificación, aumenta la diversidad de finalidades relacionadas con dicho uso y su contribución a la regulación del quehacer individual y de grupo. Sin embargo, la elaboración de la planificación por escrito y la introducción sistemática de modificaciones sigue viéndose como una carga demasiado importante por una parte importante de los alumnos.

### **Los documentos de planificación elaborados por los estudiantes**

El análisis sobre los planes de trabajo desarrollados en grupo en los tres bloques de la experiencia complementa las informaciones sobre el uso real de la planificación que realizan los estudiantes. En primer lugar, en relación con las planificaciones iniciales de los estudiantes, cabe señalar que los grupos las diseñan con distinto nivel de detalle si bien todos ellos toman como referencia los modelos que el profesorado les proporciona y explica al inicio de la experiencia. Mayoritariamente, se organizan las planificaciones bien en formato de tabla o cuadro, o bien en formato de texto narrativo; en ambos casos, en los documentos se identifican distintas marcas (colores, tipos de letra, subrayados, etc.) para señalar tareas y personas responsables de desarrollarlas. Las planificaciones explicitan, aún con distinto grado de detalle, las condiciones de trabajo de cada uno de los miembros y del grupo en su conjunto, así como las responsabilidades en relación con los roles que desempeñarán para la gestión del trabajo como resultado de las orientaciones del profesorado al respecto (secretario, dinamizador, portavoz del grupo).

En segundo lugar, destacar que es posible identificar una evolución a lo largo de los bloques temáticos en relación con los siguientes aspectos de la planificación:

- A medida que los estudiantes tienen mayor experiencia, *aumenta el detalle en la elaboración de la planificación escrita*, tanto en lo relativo a la regulación individual como de grupo. Además, los estudiantes que optan por seguir planificando por escrito en el último bloque elaboran planes con un alto nivel de detalle puesto que entienden que ello revierte en una mayor utilidad de la planificación para la gestión del aprendizaje;
- A medida que los estudiantes tienen mayor experiencia, *disminuyen las modificaciones en la planificación* a lo largo del bloque. Ello puede deberse a que, en general, los estudiantes elaboran una planificación más ajustada desde el inicio;

- *La realización de la planificación escrita disminuye* en el último bloque comparada con los anteriores. Como se recordará, en este bloque la realización de la planificación no estaba pautada con la finalidad de identificar si los alumnos habían integrado su uso en sus hábitos de trabajo.

En conjunto, se puede apreciar una mejora en los recursos de autorregulación en relación con la planificación inicial de los grupos a partir de la participación de los estudiantes en la experiencia de innovación docente. En efecto, los resultados muestran que los alumnos valoran la planificación como instrumento de regulación individual y conjunta, y que conocen mejor los elementos, las funciones y los usos relevantes de la planificación. Sin embargo, una parte importante de los estudiantes declara que considera necesaria la elaboración y el uso de la planificación para regular el aprendizaje, aunque debería ser menos formal, y otra parte aboga por una planificación no escrita (muy probablemente, oral y compartida), ya que su elaboración y seguimiento escrito les sigue pareciendo muy laborioso.

En consecuencia, a la luz de estos resultados obtenidos es posible afirmar que la influencia educativa centrada en la planificación mejora la gestión del aprendizaje por parte del alumno, pero que deben ensayarse formas de planificación menos costosas para que dicha práctica sea interiorizada por el alumno. La enseñanza de la autorregulación centrada en la planificación ha de plantearse como una intervención educativa a medio plazo, ya que necesita de la intervención y el apoyo sostenido del profesor durante un período de tiempo prolongado.

## **Resultados de la valoración de la enseñanza de la autorregulación por parte del profesorado**

Los profesores coincidieron en subrayar la relevancia de la implementación, en el diseño y desarrollo del proceso, de algunos elementos por su especial contribución a la reflexión de los alumnos sobre los procesos de autorregulación en la doble dimensión de trabajo individual y de trabajo en grupo, y por la mejora de los recursos docentes disponibles para orientar a los alumnos en su proceso de autorregulación: los elementos dirigidos a favorecer el aprendizaje de la gestión autónoma del propio aprendizaje, los elementos dirigidos a favorecer el aprendizaje y autorregulación del trabajo colaborativo y el uso de las TIC para mediar en la actividad de ajuste de la ayuda del profesor al alumno en apoyo a la autorregulación (véase Tabla IV).

TABLA IV. Recursos de ayuda educativa a la autorregulación individual y de grupo mejor valorados por el profesorado

Estructura del bloque	Recursos usados con TIC y presentados en clase presencial	Funcionalidad de la actividad conjunta profesor/alumnos para fomentar autonomía en la autorregulación individual y de grupo
Inicio del bloque	-Guía didáctica -Agenda de trabajo  -Documento sobre «Trabajo en grupo colaborativo»	-Compartir objetivos y orientar y planificar el esfuerzo individual y de grupo colaborativo.  -Representarse las características del trabajo colaborativo.
Durante el bloque	-Actividades de aprendizaje -Pautas de lectura  -Ficha de trabajo semanal en grupo -Foro de trabajo en grupo	-Pautar y guiar la realización de la actividad de enseñanza y aprendizaje  -Establecer elementos de control, regulación y ajuste en la gestión del trabajo individual y colaborativo -Traspasar progresivamente la gestión y el control -Comunicación afectiva y de ayuda mutua
Final del bloque	-Cuestionarios de autoevaluación individual y de grupo. -Devolución de los resultados de la evaluación	-Reflexionar sobre la gestión del control del aprendizaje individual y de grupo, y sobre mantener el control logrado -Comprender errores de gestión y subsanarlos

En relación con los *Elementos de apoyo al uso del trabajo colaborativo como instrumento de ayuda educativa al aprendizaje y a la autorregulación del aprendizaje*, el profesorado valoró positivamente la aportación de la experiencia de innovación a una mayor reflexión de los estudiantes sobre la regulación conjunta y sobre la ayuda de los compañeros al aprendizaje de la regulación individual. En el diseño de la propuesta docente, se tomaron decisiones sobre qué formas de organización de la actividad conjunta entre profesor y alumnos y entre alumnos podían proporcionar las condiciones deseadas para realizar todos esos aprendizajes. Las decisiones más valoradas por el profesorado son, por una parte, la selección y combinación de distintas formas de organización de la actividad conjunta de trabajo en grupo –con y sin ayuda del profesor, en diferentes espacios y momentos del proceso de enseñanza y aprendizaje, y en sesiones presenciales y virtuales– que permitieron un seguimiento por parte del profesorado y la posibilidad de poder intervenir, en consecuencia, de manera más ajustada a las necesidades del grupo a lo largo de todo el proceso. Y, en segundo lugar, los profesores valoraron que el grupo se constituyó en un espacio de apoyo mutuo entre los alumnos que contribuyó a mantener la constancia y el esfuerzo requeridos para mejorar las competencias de autorregulación y les permitió aprender a asumir y resolver los problemas de gestión conjunta del proceso, señalando el interés por conocer mejor cómo diseñar e intervenir en la construcción de los procesos socioafectivos y relacionales del grupo por su contribución al aprendizaje.

En relación con el uso de las TIC que se integran como *mediadoras de la ayuda ajustada del profesor a la autorregulación individual y de grupo*, el profesorado valoró positivamente el uso de la plataforma Moodle. Del conjunto de usos de las TIC realizado en la asignatura, de acuerdo con la clasificación propuesta por Coll (2004), los profesores destacaron que, en la experiencia de innovación, contribuyeron más directamente a la enseñanza y el aprendizaje de la autorregulación del trabajo individual y de grupo los siguientes:

- *Las TIC como apoyo a la presentación de la propuesta docente y a la explicación de los contenidos por los profesores* mediante documentos, esquemas y gráficos. Además, se integraron en la plataforma Moodle otros documentos dirigidos a favorecer la ampliación, consulta y profundización de las explicaciones del profesor.
- *Las TIC como mediadoras de la interacción entre los alumnos y los contenidos*. Las TIC se usaron como instrumentos cognitivos a disposición de los participantes al permitirles abordar las actividades de enseñanza y aprendizaje sirviéndose de pautas de lectura, de guías de trabajo, recursos para la elaboración de mapas conceptuales (*cmaptools*) y para generar productos escritos de forma colectiva (*wiki*). En definitiva, se usaron para guiar el proceso cognitivo y metacognitivo de los alumnos para apropiarse del contenido.
- *Las TIC como apoyo al seguimiento de la participación y las actuaciones de los alumnos individualmente y en grupo*. El uso de las posibilidades tecnológicas que ofrece la plataforma Moodle para registrar la actividad desplegada por los alumnos sirvió a los profesores para conocer las características de la presencia del alumno en el aula virtual (continuada y regular o discontinua y puntual) y poder actuar para apoyar y mejorar, si cabe, su implicación.
- *Las TIC como apoyo a la comunicación entre profesor y alumnos y el seguimiento y control por el profesor del proceso de aprendizaje individual y de grupo de los alumnos*. Los profesores valoraron positivamente el uso de determinados instrumentos de la plataforma, como la tutoría online y el forum de trabajo en pequeño grupo, para el seguimiento y control del trabajo del alumno individual y de grupo, y para ajustar la ayuda a las necesidades que se pusieron de manifiesto en esos espacios.
- *Las TIC como apoyo a la comunicación entre los alumnos integrantes de un grupo de trabajo y al desarrollo del trabajo colaborativo*. Los profesores valoraron que el fórum de pequeño grupo y el espacio wiki les permitieron no únicamente comunicarse, sino también identificar los problemas que surgen

en los procesos de representación de la tarea y de gestión de su desarrollo, y realizar aportaciones para guiar un trabajo que necesita de la colaboración y participación de todos para poder ser culminado con éxito.

## A modo de conclusión

A partir de los resultados de la innovación y de los datos recogidos sobre el trabajo de profesores y alumnos a lo largo de la misma, podemos concluir que:

- La enseñanza y el aprendizaje de la autorregulación individual y de grupo se puede y se debe planificar y desarrollar de forma incardinada en el proceso de enseñanza y aprendizaje;
- La organización de apoyos diferentes en diferentes momentos del proceso ayuda a identificar las necesidades del alumno en el aprendizaje de la autorregulación y a decidir tipos y grados de ayuda educativa diferentes en cada momento;
- El avance del alumno en la progresiva adquisición de la responsabilidad y la autonomía en la realización de aprendizajes significativos y con sentido se da, si éste cuenta con las ayudas necesarias. En este sentido la elaboración y uso de la planificación se ha revelado como un recurso importante de dicha adquisición;
- La enseñanza de la autorregulación de tareas realizadas en grupo supone la enseñanza y el aprendizaje de otras competencias, como las propias de la colaboración en el trabajo en equipo;
- Los profesores que trabajan colaborativamente para elaborar propuestas instruccionales innovadoras basadas en las TIC señalan que ello les supone aceptar nuevos roles y responsabilidades docentes, y que, en conjunto, les exige:
  - dedicar tiempo y esfuerzo a la reflexión, la elaboración, el desarrollo y a la evaluación conjunta de una propuesta docente compleja
  - tratar de compartir principios y criterios para una enseñanza de calidad centrada en el trabajo del alumno y dirigida al desarrollo de las competencias de éste último, en particular las de autorregulación
  - decidir usos de la tecnología relevantes, en especial aquéllos que favorecen la orientación y el seguimiento de la actividad de aprendizaje constructiva del alumno y el ajuste de la ayuda educativa a sus necesidades.

En conclusión, en este trabajo hemos tratado con mayor detalle la enseñanza y el aprendizaje de la competencia de autorregulación en el marco de una propuesta de innovación compleja. A lo largo del mismo, se ha podido comprobar la necesidad de esa enseñanza para que dicho aprendizaje se dé. Para lograrlo, se ha defendido la necesidad de presentar el aprendizaje de la autorregulación íntimamente incardinado en el proceso mismo de enseñanza y aprendizaje y se ha subrayado la exigencia de diseñar tipos y formas diversas de ayuda educativa. El profesorado implicado, que tiene experiencia en diseño y desarrollo de proyectos de innovación docente, valora positivamente los resultados obtenidos en el aprendizaje de las competencias de autorregulación y los recursos docentes diseñados para promover su enseñanza. En definitiva, los resultados permiten argumentar que colaborar para elaborar e implementar una propuesta de innovación, aún siendo un proceso costoso y complejo, resulta útil al profesorado para gestionar de modo eficaz la enseñanza, dotándola de las bases necesarias para que devenga una práctica de calidad.

## Referencias bibliográficas

- ARVAJA M., SALOVAARA, H., HÄKKINEN, P. Y JÄRVELÄ, S. (2007). Combining individual and group-level perspectives for studying collaborative knowledge construction in context. *Learning and Instruction*, 17 (4) 448-459.
- COLL, C. (2001). Constructivismo y Educación: la concepción constructivista de la enseñanza y el aprendizaje. En C. COLL, J. PALACIOS Y A. MARCHESI (comps), *Desarrollo Psicológico y Educación. Vol. 2. Psicología de la Educación Escolar* (157-186). Madrid: Alianza Editorial.
- (2004). Psicología de la educación y prácticas educativas mediadas por las tecnologías de la información y de la comunicación: una mirada constructivista. *Sinéctica*, 25, 1-24.
- COLL, C., MAURI, T. Y ONRUBIA, J. (2006). Análisis y resolución de casos-problema mediante el aprendizaje colaborativo. En A. BADÍA (coord.). *Enseñanza y aprendizaje con TIC en la educación superior* [monográfico en línea]. *Revista de Universidad y Sociedad del Conocimiento (RUSC)*. Vol. 3, n.º 2. UOC. Consultado el 20 de enero de 2007, de <http://www.uoc.edu/rusc/3/2/dt/esp/>

- ESTEVE ZARAZAGA, J. M. ET AL. (2006). Comentarios a los Informes EURYDICE y OCDE sobre la cuestión docente. *Revista de Educación*, 340, 19-86.
- JONASSEN, D. H. Y HERNÁNDEZ-SERRANO, J. (2002). Case based Reasoning and Instructional Design: Stories to Support Problem Solving. *Educational Technology Research and Development*, 50 (2), 65-77.
- MAURI, T., ONRUBIA, J., COLL, C. Y COLOMINA, R. (2005). La calidad de los contenidos educativos reutilizables: diseño, usabilidad y prácticas de uso. *RED. Revista de Educación a Distancia. Monográfico II*. Consultado el 20 de enero de 2007, de <http://www.wa/ead/red/M2>
- Onrubia, J. (2005). Aprender en entornos virtuales de enseñanza y aprendizaje: actividad conjunta, ayuda pedagógica y construcción del conocimiento. *RED. Revista de Educación a Distancia. Monográfico II*. Consultado el 20 de enero de 2007, de <http://www.wa/ead/red/M2>
- PERRENOUD, P. (2004). *Desarrollar la práctica reflexiva en el oficio de enseñar. Profesionalización y razón pedagógica*. Barcelona: Graó.
- PERRY, N. E., VAANDEKAMP, K. O., MERCER, L. K. Y NORDBY, C. J. (2002). Investigating Teacher-Student Interactions That Foster Self-Regulated Learning. *Educational Psychologist*, 37, (1) 5-15.
- PINTRICH, P. R. Y ZUSHO, A. (2002). The development of academic self-regulation: The role of cognitive and motivational factors. En A. WIGFIELD Y J. S. ECCLES (eds.). *Development of achievement motivation* (pp. 249-284). San Diego: Academic Press.
- POZO, J. I. Y MONEREO, C. (2002). Introducción. Un currículo para aprender. Profesores, alumnos y contenidos ante el aprendizaje estratégico. En J. I. POZO Y C. MONEREO (coords.). *El aprendizaje estratégico. Enseñar a aprender desde el currículo* (pp. 11-25). Madrid: Aula XXI/ Santillana.
- SCHUNK, D. H. Y ZIMMERMAN, B. J. (2003). Social origins of self-regulatory competence. *Educational Psychologist*, 32, 195-208.
- VYGOTSKI, L. S. (1925). Soznaniye kak problema psikhologii povenedija. En K.N. Kornilov (Ed.) *Psikhologiya i marksizm*. Moscou: Gosizdat.
- ZEIDNER, M., BOEKAERTS, M. Y PINTRICH, P. R. (2000). Self-regulation. Directions and challenges for future research. En M. BOEKAERTS, P. R. PINTRICH Y M. ZEIDNER (eds.). *Handbook of Self-Regulation* (pp. 749-768). San Diego: Academic Press.

**Dirección de contacto:** Teresa Mauri Majós. Universidad de Barcelona. Departamento de Psicología Evolutiva y de la Educación. Passeig de la Vall d'Hebron, 171. 08035 Barcelona. E-mail: [teresamauri@ub.es](mailto:teresamauri@ub.es)



# Universitarios de ciencias ante la docencia en Educación Secundaria como expectativa profesional

## University students of sciences and secondary teaching as a professional prospect

Francisco Córdoba Alcaide

Rosario Ortega Ruiz

*Universidad de Córdoba. Facultad de Ciencias de la Educación. Departamento de Psicología. Córdoba, España*

Alfonso Pontes Pedrajas

*Universidad de Córdoba. Departamento de Física Aplicada. Córdoba, España*

### Resumen

Este estudio descriptivo a través de encuestas analiza el interés por la enseñanza en Educación Secundaria como salida profesional entre universitarios del Área de Ciencias Experimentales y los motivos por los que podrían elegir esta salida. El recorrido por la literatura que se expone alerta de la necesidad de llevar a cabo un análisis de dichas motivaciones previas así como de la idoneidad de tenerlas en cuenta en un proceso selectivo de aspirantes a docentes. Se ha realizado encuestando a 315 universitarios de las carreras de Física, Química y Biología acerca de sus percepciones sobre la profesión docente y su interés dentro de un rango de preferencia. Observamos que una elevada proporción de sujetos contempla la docencia en Educación Secundaria como salida profesional posible durante los años de sus licenciaturas pero no como la preferente, lo cual iría en la línea de trabajos que aseguran que el alumnado que se dedica a la enseñanza en Educación Secundaria en su gran mayoría tenía otras expectativas profesionales y que las dificultades del mundo laboral les han hecho reconsiderar su postura. Además, aparece como elemento innovador el relacionar la preferencia o no por la profesión docente con los motivos por los que se elegiría y revela cómo los sujetos que proyectan la docencia en posición preferente expresan motivos relacionados con su interés por

determinadas facetas de la labor docente; cuando la proyectan en posición intermedia, los motivos se relacionan con estereotipos acerca de los beneficios de la profesión, y, por último, si se visualiza en posición final, los motivos tienen que ver con el hecho de no desechar ninguna posibilidad *a priori*.

*Palabras clave:* interés por la docencia, Educación Secundaria, formación inicial del profesorado, motivaciones previas, expectativas profesionales.

### **Abstract**

This descriptive study via surveys analyses the interest in secondary teaching as a professional option for university students of Experimental Sciences and the reasons for them to choose it. We present a review of the literature which alerts us to the necessity of carrying out an analysis of these previous motivations as well as of the suitability to take them into account in a selective process of individuals who want to be teachers. 315 university students enrolled in Physics, Chemistry and Biology have been surveyed about their perceptions of teaching in Secondary Education as a professional option within a range of preferences. We observed that a high proportion of students from the sample perceive teaching in Secondary Education as a possible professional option during the years in which they are doing their degrees, but not as the preferred one. This observation supports the line of investigations that assure that most students who work as teachers in Secondary Education are students who had other professional expectations, but that the difficulties to find a job, have made them reconsider their position. In addition, an innovative element appears via relating the preference or non-preference of the teacher profession to the reasons by which it would be chosen, and it reveals how the subjects that project Secondary teaching as their preferred position express reasons related to their interest in specific facets of the educational work. When the subjects place it in intermediate position, the reasons are related to stereotypes about the benefits of the profession; and, finally, if teaching is seen in final position, the reasons have to do with the belief of not rejecting any possibility, in principle.

*Key Words:* interest in teaching, Secondary Education, Initial Teacher Training, previous motivations, profesional expectative.

## Introducción

Al surgir en las dos últimas décadas nuevas leyes que han modificado notablemente la estructura de la Educación Secundaria, las diferentes administraciones educativas han intentado desarrollar un modelo de formación docente más adecuado que el curso para la obtención del Certificado de Aptitud Pedagógica (CAP) proponiendo diversas alternativas como el Curso de Cualificación Pedagógica (CCP), que sólo llegó a implantarse de forma experimental en algunas universidades del país durante los años noventa, o el curso para la obtención del Título de Especialización Didáctica (TED), que apareció en un decreto ley del año 2004 y no llegó a implantarse porque fue derogado unos meses después de su publicación.

La crítica más consistente al CAP lo considera como «un añadido pedagógico», en el que se pretende construir en pocos meses una *identidad profesional* suplementaria para que las personas que finalizan sus estudios universitarios con una formación muy general se conviertan en docentes (Esteve, 1997). Por otro lado, la formación académica, que no recibe apenas crítica alguna y que, por el contrario, es altamente valorada, se realiza en el marco de los estudios universitarios ordinarios (estudios de grado, a partir del 2010) al margen de cualquier formación psicoeducativa. Este modelo de formación no deja de ser una yuxtaposición a la identidad del candidato que ya queda establecida como científico, o simplemente como licenciado en la materia de su formación académica (Gimeno, 1988).

En este sentido, adquiere la máxima importancia la nueva Ley Orgánica de Educación (LOE), en la que el título III se dedica expresamente al profesorado en general y, en concreto, al de la etapa de Educación Secundaria y a su protagonismo en el proceso de reforma. En él se hace referencia y se presta una atención prioritaria a su formación inicial y permanente, cuyas modificaciones deben llevarse a cabo en los próximos años, en el contexto del Espacio Europeo de Educación Superior (EEES), para dar respuesta a las necesidades y a las nuevas demandas que recibe el sistema educativo. El MEC (2006) ha presentado recientemente una propuesta de título universitario oficial para enseñanzas de posgrado bajo la denominación de Máster en Formación de Profesorado de Educación Secundaria y que pretende acercarse a la realidad del momento de acuerdo a lo establecido en la LOE.

Si partimos de la idea de que la formación del profesorado es un proceso largo y continuo, parece lógico considerar que en el período de formación inicial no es posible proporcionar al futuro docente todo el conocimiento profesional que necesitará en el complejo ejercicio de sus futuras tareas profesionales, ya que hay

situaciones que sólo llegan a resolverse y a conocerse en profundidad gracias a la experiencia y a la autoevaluación de las competencias adquiridas durante el desarrollo profesional (Cano, 2005; Moreno, 2006). Pero ello no significa que haya que obviar por completo este período de formación inicial, como se ha venido haciendo desde que fue regulado legalmente en los años 70, para una sociedad con unas necesidades muy diferentes a las de la actualidad (Córdoba y Ortega, 2006; Córdoba, Ortega y Monks, 2006).

Por nuestra parte, a la hora de delimitar el período de formación inicial tomaremos la propuesta de Michavilla (2004) sobre la formación del docente universitario y que extrapolamos a la formación del docente de Educación Secundaria. Este autor diferencia entre tres períodos claves dentro de la formación: el período de la formación previa, que incluiría aquella formación que lleva a cabo el sujeto antes de tomar la decisión de dedicarse a una profesión determinada; el período de formación inicial, como aquél que se dedica a proporcionar competencias profesionales específicas como prerrequisito al desempeño de la profesión, y el período de la formación permanente, dedicado a que el sujeto se recicle y adquiera aquellas competencias que demande la sociedad en un determinado momento.

## **El estudio del interés por la enseñanza como salida profesional**

Para obtener datos sobre el interés o la motivación previa con respecto a la enseñanza como salida profesional, se han realizado estudios en los que se preguntaba a los docentes en ejercicio cuáles fueron sus motivos para dedicarse a la enseñanza (González Blasco y González-Anleo, 1993; Esteve, 1997; Sepúlveda y Rivas, 2000) y otros en los que se pregunta a los futuros docentes por sus preferencias, fundamentalmente a los alumnos del curso del CAP (Campanario, 1998; García y Martínez, 2001), pero las expectativas con respecto a la docencia como salida profesional en aspirantes a la FIPS (es decir, durante el período de formación previa) es un aspecto que ha recibido mucha menos atención (Córdoba, Ortega y Pontes, 2006; Córdoba, Ortega y Romera, 2006).

Esteve (1997) ofrece una revisión de algunos de los investigadores que se han dedicado al estudio de las motivaciones de ingreso a la docencia: Amiel (1980), Berger (1976), León (1980), Varela y Ortega (1984), Meirieu (1991), entre otros, los cuales concluyen destacando una concepción ideal de ayuda y de relación interpersonal individual como motivo central del ingreso en la profesión docente. En esta revisión,

dicho autor incluye trabajos en los que se añade como característica de la elección profesional la seguridad que ofrece la situación de funcionario (Berger, 1976).

Los estudios revisados por Esteve (1997) señalan la necesidad de indagar acerca de las motivaciones que llevan a los universitarios, a los aspirantes a la FIPS, a proyectar o no la docencia en Educación Secundaria como una de sus salidas profesionales y a considerar dichas motivaciones previas incluso como un factor más que será preciso tener en cuenta en el proceso de selección del profesorado de esta etapa educativa. Como consecuencia de los trabajos revisados, el citado autor asegura que el alumnado que se dedica en España a la enseñanza en Educación Secundaria es, en su gran mayoría, alumnado que tenía otras expectativas profesionales y al que, sin embargo, las dificultades del mundo laboral le ha hecho reconsiderar su postura y dedicarse a la enseñanza como último recurso (Esteve, 1997).

En este sentido, si analizamos un atractivo trabajo que relata las biografías profesionales de seis profesores de Educación Secundaria, nos encontramos con que en cinco de ellas se admite que la enseñanza no era contemplada como preferencia personal. Cinco de estos profesionales admiten retrospectivamente que se dedicaron a la enseñanza más por obligación que por vocación y justifican el hecho de haber acabado siendo docentes aludiendo a situaciones como la escasez de salidas profesionales. Sólo en uno de los casos se admite que la enseñanza era la salida profesional preferente (Sepúlveda y Rivas, 2000).

Por el contrario, González Blasco y González-Anleo (1993) señalaban, en un estudio previo, que, entre el profesorado en ejercicio de niveles no universitarios, la vocación como rasgo de la profesionalidad es reconocida por casi la totalidad del profesorado de su muestra en el ámbito nacional. La vocación aparece como el motivo más importante para dedicarse a la enseñanza, junto a otras razones como el humanismo de la profesión o la facilidad o conveniencia de la carrera de profesor. Según estos autores, las investigaciones sobre los motivos para seguir la carrera de maestro o de profesor no han llegado a conclusiones definitivas pero se destaca el factor del idealismo.

En las investigaciones efectuadas con alumnado que realiza el curso del CAP, cuando se le pregunta por el interés que le ha llevado a seguir dicho curso, paradójicamente el interés por la enseñanza y la intención de ejercer como docentes aparece devaluada. En un trabajo reciente (García y Martínez, 2001), se destaca que el hecho de que el CAP sea el requisito necesario para acceder al cuerpo de profesores de Educación Secundaria es lo que suscita el interés de un 88,6% del alumnado de la muestra para realizarlo. Le sigue en importancia la necesidad de aprender aspectos relacionados con la enseñanza (47,6%) y, en último lugar, el interés por la enseñanza (19%).

## Objetivos del estudio

Como se puede deducir a partir de la anterior revisión bibliográfica, el estudio prospectivo de la motivación e interés por la enseñanza como salida profesional ha sido objeto de menor atención por los investigadores e investigadoras al considerarse tradicionalmente la formación previa, de carácter academicista, como algo desligado del proceso de formación del docente de Educación Secundaria. En nuestra opinión, la etapa de formación universitaria previa ofrece un contexto útil e interesante para explorar las intenciones y expectativas de los estudiantes de Ciencias sobre la profesión docente. Por ello, en este trabajo pretendemos proporcionar datos empíricos acerca de los motivos que, en la actualidad, acompañan a la intención de los universitarios (posibles aspirantes a la FIPS) de elegir la enseñanza en Educación Secundaria como una de sus salidas profesionales.

En concreto, los objetivos específicos de este estudio han sido los siguientes:

- Conocer la posición que ocupa la enseñanza entre las expectativas profesionales de los universitarios de Ciencias.
- Describir los motivos a los que aluden los universitarios de la muestra en su intención de elegir la enseñanza en Educación Secundaria como una de sus salidas profesionales.
- Relacionar la posición en la que los universitarios de la muestra incluirían la enseñanza como salida profesional, con los motivos que les llevarían a elegirla.

## Metodología

### Diseño de la investigación

Para la realización de este trabajo, hemos utilizado una metodología descriptiva a través de encuestas (Cardona, 2002) y un diseño transversal (Buendía, Colás y Hernández, 1998). Esta metodología de investigación posibilita recabar información significativa acerca del hecho que estamos investigando, ya que nos ha permitido conocer las intenciones profesionales de los sujetos de la muestra, así como los motivos que los llevarían a elegir o no la enseñanza entre sus salidas profesionales.

## Muestra

Para la selección de los participantes en la muestra, realizamos un muestreo probabilístico aleatorio estratificado y seleccionamos tres licenciaturas del área Científico-Tecnológica (Biología, Física y Química). El criterio que se tuvo en cuenta a la hora de seleccionar las licenciaturas fue que se tratara de estudios tradicionalmente ligados a áreas curriculares en Educación Secundaria. Además, nos interesaba comparar las respuestas de los participantes de la muestra al comienzo de sus estudios universitarios (alumnado de primer curso), en el ecuador de sus estudios (tercer curso) y en los últimos cursos (cuarto y quinto curso).

Hemos seleccionado una muestra formada por 315 estudiantes (121 varones y 194 mujeres) de tres licenciaturas del área de Ciencias Experimentales (Biología, Física y Química) para su comparación.

La distribución de los cuestionarios entre los diferentes grupos fue realizada mediante afijación proporcional (Buendía, Colás y Hernández, 1998). En la Tabla I, presentamos las características básicas de la muestra de estudiantes con la que hemos trabajado en este estudio exploratorio, en función de la licenciatura y del curso de las personas encuestadas.

TABLA I. Características de la muestra del estudio

LICENCIATURA	PRIMER CICLO		SEGUNDO CICLO		TOTAL
	1º Curso	3º Curso	4º Curso	5º Curso	
Biología	80	51	--	35	166
Física	8	13	23	--	44
Química	47	27	31	--	105
<b>TOTAL</b>	135	91	54	35	315

## Instrumento de investigación

Para alcanzar los objetivos de la investigación, hemos elaborado un cuestionario que permite recoger las respuestas del alumnado de la muestra a los problemas de investigación planteados (Córdoba, Ortega y Rodríguez, 2005). Para el diseño de dicho instrumento, se realizó una primera versión del cuestionario y se llevó a cabo una primera validación interjueces a partir de las aportaciones realizadas por cuatro profesores doctores de la Universidad de Córdoba, que pueden ser considerados expertos en investigación educativa o en formación del profesorado. Tales expertos

fueron entrevistados por uno de los autores de este trabajo para recoger todas aquellas modificaciones y aportaciones que considerasen necesarias, llegando así a un primer borrador que posteriormente aplicamos, en forma de experiencia piloto, con 20 estudiantes de la muestra. Los comentarios de estos estudiantes nos sirvieron para realizar algunas modificaciones en la estructura del cuestionario inicial hasta llegar a obtener la versión definitiva del cuestionario que se muestra en el Anexo I.

## **Procedimiento de recogida y análisis de los datos**

La recogida de los datos se llevó a cabo al comienzo de una clase de los respectivos cursos encuestados, presentando el cuestionario a los participantes (con el permiso previo del profesorado de asignaturas obligatorias o troncales) mediante una breve explicación de los fines del estudio y haciendo hincapié siempre en la voluntariedad del mismo. La tasa de participación del alumnado ha sido del 100% en todas las clases y el tiempo medio empleado en responder al cuestionario ha sido de 15 minutos.

La codificación, sistematización y análisis de los datos cualitativos y cuantitativos se realizó mediante el paquete estadístico SPSS+ (versión 12.0). Previamente a la codificación de los datos cualitativos referidos a la pregunta abierta: *¿Ha incluido la enseñanza en Educación Secundaria entre las salidas profesionales de su carrera? ¿Por qué?*, se elaboró un sistema de categorías a posteriori. De esta manera, hemos ido definiendo las categorías utilizadas a medida que examinábamos los datos, siguiendo un procedimiento inductivo (Strauss, 1987).

El criterio fundamental que seguimos para la extracción de las categorías de respuesta fue el temático, en función del contenido que el participante abordaba en su respuesta. El procedimiento de elaboración del sistema de categorías partió de la exploración de los sistemas ya existentes utilizados por otros investigadores en estudios similares o compatibles con nuestra investigación (Amiel, 1980; Varela y Ortega, 1984; González Blasco y González-Anleo, 1993; García y Martínez, 2001). Una vez analizados los sistemas de atribución de motivos más reconocidos en la literatura, seleccionamos aleatoriamente un 30% de los cuestionarios para que dos expertos, de forma separada, realizaran un sistema previo de categorías. Tras varias sesiones de trabajo independiente por parte de ambos jueces, se presentaron y negociaron ambos sistemas de categorías, así como las definiciones de las mismas hasta llegar al acuerdo definitivo sobre las mismas que aparece en el Cuadro I.

CUADRO I. Categorías de atribución de motivos que acompañan a la intención de elegir la profesión docente en Educación Secundaria

CATEGORÍA	DEFINICIÓN	EJEMPLO
(1) SIMPATIZANTES	Expresan interés por la enseñanza en Educación Secundaria como salida profesional que les interesa o les gusta, con argumentos referidos a determinadas facetas de la labor del docente tales como la transmisión de conocimientos o el contacto con el alumnado.	«Porque me gustaría dedicarme a la docencia» (Varón, 3º Biología).
(2) VOCACIONALES	Expresan que se sienten comprometidos con la enseñanza como salida profesional utilizando argumentos relacionados con aspectos básicos de la actividad profesional del docente y/o afirmando que la enseñanza es una profesión que ha estado siempre entre sus intereses o es vocacional. Eligen esta salida profesional porque es el objetivo principal hacia el que han orientado su vida o el desarrollo profesional de su licenciatura.	«Elegí esta carrera porque me interesaba esta salida, ya que siempre me ha gustado la educación» (Varón, 1º Química)
(3) PRAGMÁTICOS	Expresan que su interés por la enseñanza tiene que ver únicamente con ciertos estereotipos acerca de los beneficios de la profesión del docente (muchas vacaciones y buen sueldo; estabilidad económica o laboral) y con las características del sistema de acceso.	«Aprobando las oposiciones te aseguro un trabajo para toda la vida, con muchas vacaciones». (Varón, 4º de Física)
(4) DESINFORMADOS	Expresan que incluyen la enseñanza en Educación Secundaria como una de sus salidas profesionales porque no conocen otras salidas, dados los estudios universitarios que realizan.	«Porque es una de las pocas salidas profesionales que conozco de mi titulación». (Mujer, 1º Biología)
(5) ALTERNATIVOS	Eligen la enseñanza en Educación Secundaria porque no quieren desechar ninguna posibilidad a priori aunque aseguran que sólo se dedicarían a esta profesión si no logran en poco tiempo otras metas que valoran más.	«Porque es una posibilidad». (Varón, 1º Física)
(6) DESINTERESADOS	No incluyen la enseñanza en Educación Secundaria como una de sus salidas profesionales preferentes	«No me atrae la enseñanza y menos en secundaria». (Mujer, 5º Biología)
(7) OTROS	Respuestas que no se pueden categorizar de acuerdo con el sistema propuesto por la ambigüedad de las expresiones. Fundamentalmente, se trata de respuestas en las que aparecen simultáneamente dos o tres de las categorías anteriores en la misma respuesta. Este tipo de respuestas simultáneas han sido minoritarias (2,9 %).	«Es lo primero que se te viene a la mente». (Varón, 4º Química)
(8) NS/NC	No expresan ningún motivo.	

Una vez establecido el sistema de categorías que acabamos de presentar, volvimos a seleccionar aleatoriamente un 30% de los cuestionarios de la muestra, que fueron categorizados respectivamente por ambos jueces de manera independiente para la aplicación de la medida Kappa del acuerdo interjueces (Cohen, 1960). El acuerdo perfecto aparece cuando  $K=1$ , mientras que la independencia de criterio en la categorización implica que  $K=0$  (Aguilera, 2001); en la Tabla III, aparecen las categorías de atribución de motivos junto con el valor de Kappa para cada una de ellas. Como puede observarse, para la mayor parte de las categorías el acuerdo interjueces es perfecto ( $K=1$ ), excepto en las categorías «simpatizantes» ( $K=0,977$ ), «desinformados» ( $K=0,949$ ) y «alternativos» ( $K=0,934$ ), en las que se podría considerar «casi perfecto»

de acuerdo con el sistema que estableció Everitt (1992) para evaluar el grado de acuerdo ya que K se encuentra en el intervalo entre 0,81 y 1,00.

CUADRO II. Prueba Kappa de acuerdo interjueces sobre las categorías de atribución de motivos

CATEGORÍA:	K:
1 = SIMPATIZANTES	0,977
2 = VOCAACIONALES	1,000
3 = PRAGMÁTICOS	1,000
4 = DESINFORMADOS	0,949
5 = ALTERNATIVOS	0,934
6 = DESINTERESADOS	1,000
7 = OTROS	1,000
8 = NS/NC	1,000

## Resultados

### La docencia en Educación Secundaria como salida profesional y la posición que ocupa esta profesión entre las intenciones del alumnado universitario

Para hacer más significativo el análisis de los resultados de este apartado, hemos agrupado las cinco posiciones en las que los sujetos podían incluir la enseñanza en Educación Secundaria como salida profesional (véase Anexo I) en tres situaciones a las que hemos denominado respectivamente: preferente (posición primera), intermedia (posiciones segunda y tercera) y final (posiciones cuarta y quinta).

Los resultados del estudio obtenidos sobre este aspecto, que están recogidos en la Tabla II, nos muestran que un 75,2% de los universitarios de las tres carreras de Ciencias analizadas incluyen la enseñanza en Educación Secundaria entre sus posibles salidas profesionales frente al 24,8% que no la incluye.

TABLA II. Posición de la docencia en Educación Secundaria como salida profesional de los estudiantes de Ciencias

POSICIÓN	FRECUENCIA	PORCENTAJE	PORCENTAJE VÁLIDO	PORCENTAJE ACUMULADO
Preferente	35	11,1 %	11,1 %	11,1 %
Intermedia	108	34,3 %	34,3 %	45,4 %
Final	94	29,8 %	29,8 %	75,2 %
Excluye la enseñanza	78	24,8 %	24,8 %	100,0 %
Total	315	100,0 %	100,0 %	100,0 %

Al realizar un análisis más detallado de los datos anteriores, en función de la licenciatura, encontramos que el 72,3% de los sujetos en Biología, el 70,5% en Física y el 81,9% en Química ven la docencia en Educación Secundaria como posible salida.

Como podemos observar en la Tabla III, en las tres licenciaturas de Ciencias la tendencia general de los sujetos de la muestra es la de proyectar la docencia como salida profesional en las posiciones intermedia y final. En la Licenciatura de Biología, la intención de elegir la docencia se sitúa en posición intermedia en el primer curso (36,3%), en tercer curso en posición final (37,3%) y en quinto de nuevo en intermedia (45,7%). En Física, la docencia se proyecta mayoritariamente en posición intermedia (62,5%) en el primer curso y ninguno de los sujetos de la muestra se plantea la docencia como salida profesional en posición preferente o final; en tercer curso sigue siendo la posición intermedia la más frecuente (30,8%) y, en cuarto curso, las posiciones intermedia y final aparecen elegidas con el mismo porcentaje (26,1%). Química es la única licenciatura en la que en primer curso las posiciones preferente e intermedia aparecen con el mismo porcentaje (21,3%); en tercero y cuarto aparece la posición intermedia como la más elegida (48,2% y 45,2%, respectivamente).

TABLA III. Posición de la docencia en Educación Secundaria como salida profesional de los estudiantes de Ciencias

POSICIÓN	BIOLOGÍA			FÍSICA			QUÍMICA		
	1º curso	3º curso	5º curso	1º curso	3º curso	4º curso	1º curso	3º curso	4º curso
Preferente	8,8%	7,8%	8,6%	0%	15,4%	21,7%	21,3%	7,4%	6,5%
Intermedia	36,3%	21,6%	45,7%	62,5%	30,8%	26,1%	21,3%	48,2%	45,2%
Final	27,4%	37,3%	25,7%	0%	23%	26,1%	38,3%	25,9%	32,2%
Excluye la enseñanza	27,5%	33,3%	20,0%	37,5%	30,8%	26,1%	19,1%	18,5%	16,1%

## Motivos de la intención de elegir la enseñanza en Educación Secundaria como salida profesional

El análisis de los motivos que indican los sujetos de la muestra para justificar su posición con respecto a la enseñanza en Educación Secundaria como posible salida profesional ha permitido clasificarlos en las categorías que se muestran en la Tabla IV, donde podemos comprobar que, de forma general, tales estudiantes de Ciencias se sitúan principalmente como «alternativos» (20,6%) y «pragmáticos» (19,4%). Es decir, el alumnado de la muestra, de forma general, se plantea la enseñanza en Educación Secundaria ya desde la licenciatura porque no quiere desechar ninguna posibilidad a

priori, es decir, como alternativa si no encuentra el trabajo deseado; de ahí el nombre de la categoría «alternativos». Además, otro elevado porcentaje lo hace por motivos que tienen que ver con los estereotipos que rodean la profesión docente, como las amplias vacaciones, el tiempo libre, etc., siendo éste el caso de los que definíamos, en el Cuadro I, como «pragmáticos».

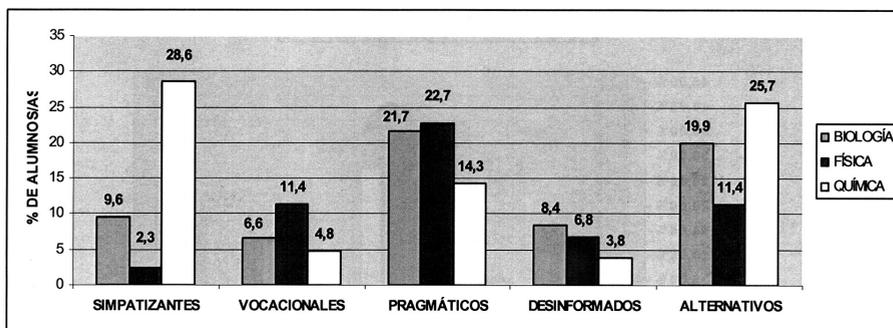
TABLA IV. Clasificación de los motivos de elección de la docencia en Educación Secundaria como salida profesional

CATEGORÍAS	FRECUENCIA	PORCENTAJE VÁLIDO	PORCENTAJE ACUMULADO
Simpatizantes	47	14,9	14,9
Vocacionales	21	6,7	21,6
Pragmáticos	61	19,4	41,0
Desinformados	21	6,7	47,6
Alternativos	65	20,6	68,3
Otros	9	2,9	71,1
NS/NC	13	4,1	75,2
Desinteresados	78	24,8	100,0
Total	315	100,0	100,0

Al describir los motivos de elección en función de la licenciatura, como se puede apreciar en la Figura I, los universitarios de Biología de la muestra que incluyen la enseñanza eligen en primer lugar respuestas propias de la categoría «pragmáticos» (21,7%) y «alternativos» (19,9%). En Física, también la categoría «pragmáticos» es la que mayores porcentajes registra (22,7%), seguida en este caso de «vocacionales» y «alternativos» con el mismo porcentaje (11,4%). Y en Química, la categoría que mayores porcentajes registra es «simpatizantes» (28,6%), seguida de «alternativos» (25,7%).

Siguiendo la tendencia general, la categoría «pragmáticos» es la que mayores porcentajes registra en el caso de las licenciaturas de Biología y Física, pero no es así en el caso de Química, en la que la categoría con mayor porcentaje es «simpatizantes». Hemos utilizado dicha categoría para referirnos al alumnado de la muestra que elige la enseñanza porque le interesa o porque le gustan determinadas facetas de la labor docente como el trato con el alumnado, la transmisión de conocimientos, etc. Si unimos este dato al anterior, es decir, al hecho de que el alumnado de primer curso de Química sea el único que tiene la intención de elegir la docencia en Educación Secundaria como salida profesional preferente, ahora nos encontramos con que en general el alumnado de esta licenciatura presenta especial simpatía por la docencia. Dado el carácter descriptivo de este trabajo, sería motivo de un análisis más completo el determinar las causas de estas diferencias así como su significatividad estadística.

FIGURA I. Motivos para elegir la profesión docente en Educación Secundaria en función de la licenciatura



Entre los motivos de inclusión en función del curso (véase Tabla V), en la Licenciatura de Biología la categoría «pragmáticos» es la que mayor porcentaje registra en el primer curso (18,8%); en tercer curso es la categoría «alternativos» (27,5%), y en quinto curso vuelven a aparecer los «pragmáticos» en primer lugar (31,4%). En Física, se elige mayoritariamente también la categoría «pragmáticos» (25,0%) en el primer curso, en el que las categorías «simpatizantes», «vocacionales» y «desinformados» no son elegidas por ninguno de los participantes; en tercer curso, las categorías «vocacionales» (23,1%) junto con la «pragmáticos» (23,1%) son las que mayores porcentajes registran y de nuevo la categoría «simpatizantes» no aparece elegida; en cuarto curso, vuelven a destacar los «pragmáticos» (21,7%). En Química, la categoría «simpatizantes» es la más elegida en primero y tercero (36,2% y 33,3%, respectivamente); sin embargo, en cuarto curso la categoría «alternativos» (35,5%) pasa a ocupar el primer lugar y la categoría «desinformados» no es elegida por ninguno/a de los participantes de este curso.

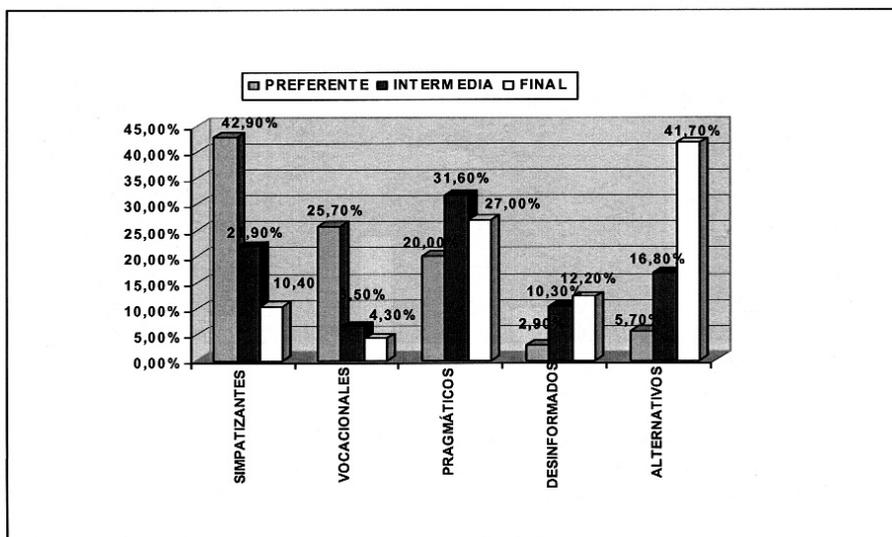
TABLA V. Atribución de los motivos en la elección de la docencia en Educación Secundaria como salida profesional en función de la licenciatura y el curso.

	BIOLOGÍA			FÍSICA			QUÍMICA		
	1º curso	3º curso	5º curso	1º curso	3º curso	4º curso	1º curso	3º curso	4º curso
SIMPATIZANTES	10,0%	9,8%	8,6%	0%	0%	4,3%	36,2%	33,3%	12,9%
VOCACIONALES	5,0%	5,9%	11,4%	0%	23,1%	8,7%	0%	3,7%	12,9%
PRAGMÁTICOS	18,8%	19,6%	31,4%	25,0%	23,1%	21,7%	8,5%	14,8%	22,6%
DESINFORMADOS	12,5%	3,9%	5,7%	0%	7,7%	8,7%	4,3%	7,4%	0%
ALTERNATIVOS	16,3%	27,5%	17,1%	12,5%	7,7%	13,0%	25,5%	14,8%	35,5%

## Relación entre la intención de elegir la enseñanza en Educación Secundaria como salida profesional y motivos de dicha intención

En este apartado, se describen, de forma general, los motivos que acompañan a la intención del alumnado participante de incluir la enseñanza en Educación Secundaria como salida profesional en una u otra posición. En la Figura II, se muestra la dependencia que existe entre los motivos de elección y la posición (preferente intermedia o final) en que se proyecta dicha elección.

FIGURA II. Motivos para elegir la profesión docente en función de la posición de elección



Puede apreciarse que, para el alumnado de la muestra que incluye la enseñanza en Educación Secundaria en posición preferente, la categoría con el porcentaje más alto es la de «simpatizantes» (42,8%) seguida de «vocacionales» (25,7%). Para los que incluyen la enseñanza en Educación Secundaria en posición intermedia, las categorías más elegidas son «pragmáticos» (31,6%), seguida de «simpatizantes» (21,9%). Finalmente, el alumnado que elige la enseñanza en la posición final elige la categoría «alternativos» (41,7%) en primer lugar y «pragmáticos» (27,0%) en segundo.

## Análisis de resultados

Los resultados empíricos de este estudio descriptivo, acerca de los motivos que acompañan a la intención del alumnado universitario de ciencias de elegir la profesión docente en la Educación Secundaria, nos parecen útiles y necesarios para abordar cualquier proyecto de transformación del modelo actual de formación pedagógica inicial del profesorado de esta etapa educativa.

En este contexto, debemos resaltar, en primer lugar, que la mayoría del alumnado de carreras de ciencias que ha participado en este estudio proyecta, es decir, manifiesta la intención de elegir la enseñanza en Educación Secundaria entre sus salidas profesionales en algún grado o, como hemos llamado aquí, en alguna posición.

Como se ha indicado anteriormente, existe una línea de trabajos relacionados con la FIPS que aseguran que el alumnado que se dedica a la enseñanza en Educación Secundaria es, en su gran mayoría, alumnado que tenía otras expectativas profesionales y a los que, sin embargo, las dificultades del mundo laboral les han hecho reconsiderar su postura y dedicarse a la enseñanza como último recurso (Esteve, 1997; Sepúlveda y Rivas, 2000, García y Martínez, 2001). En este sentido, en las tres licenciaturas estudiadas, la tendencia general de los sujetos de la muestra es la de elegir la docencia como salida profesional en las posiciones intermedia y final. La posición preferente aparece como la menos elegida en la mayoría de los casos. Por lo general, los universitarios de ciencias de la muestra contemplan la docencia en Educación Secundaria como una posible salida profesional pero no como la preferente, lo cual iría en la línea de los trabajos citados.

Estos resultados contrastan con el elevado número de alumnos que, una vez que terminan alguna de estas licenciaturas, deciden matricularse en el curso para la obtención del CAP sin que, como acabamos de comprobar, la docencia en Educación Secundaria sea algo preferente para ellos y ellas. La facilidad que entraña la realización de este curso en la mayoría de las universidades españolas, su escasa duración y la obtención del título de manera ya ni siquiera presencial podrían estar en la base de esta decisión que toman muchos estudiantes y que supone la realización del CAP «por si acaso puede servir para algo en el futuro».

Si reflexionamos sobre los motivos que acompañan a la intención de los sujetos de la muestra de incluir la docencia en Educación Secundaria entre sus salidas profesionales, debemos resaltar que, de forma general, estos sujetos se identifican principalmente con las categorías «pragmáticos» y «alternativos». Es decir, parecen inclinarse por motivos que tienen que ver con ciertos estereotipos de la profesión relacionados con el

tiempo libre, el dinero, las vacaciones, etc., en el caso de la categoría «pragmáticos», en la que encontramos comentarios del tipo: «He incluido la docencia en Secundaria porque los profesores tienen mucho tiempo libre y muchas vacaciones», «porque se trata de una profesión que da estabilidad económica y laboral», etc. Y, también, porque no quieren desechar ninguna posibilidad a priori, aunque aseguran que solo se dedicarían a esta profesión si no lograran en poco tiempo otras metas que valoran más. Éste es el caso de los «alternativos», que realizan afirmaciones como: «Elijo la enseñanza porque es una salida más», «porque no quiero desechar ninguna posibilidad», etc.

Al analizar estas motivaciones hacia la docencia en función de las licenciaturas, observamos que la tendencia general, señalada anteriormente, aparece más pronunciada en el caso de Biología y Física. En Química, sin embargo, el alumnado de la muestra ofrece respuestas propias de la categoría «simpatizantes», que hemos utilizado para representar al alumnado de la muestra que elige la enseñanza porque le interesa o le gusta determinadas facetas de la labor docente como puede ser el trato con el alumnado, la transmisión de conocimientos, etc. Este alumnado realiza afirmaciones del tipo: «Me gusta la enseñanza», «me gustaría enseñar porque me encanta el trato con los adolescentes», etc. Es decir, el alumnado de Química manifiesta cierta simpatía por la docencia en Educación Secundaria como profesión que les gusta o atrae, lo cual será objeto de especial interés en próximas investigaciones.

Continuando con el análisis de la atribución de motivos, pero considerando además la posición en la que cada persona de la muestra tiene la intención de incluir la docencia en Educación Secundaria como salida profesional, encontramos algunas diferencias. Cuando los sujetos proyectan la docencia en Educación Secundaria en posición preferente, expresan motivos propios de la categoría «simpatizantes»; cuando la proyectan en posición intermedia, los motivos comienzan a estar más relacionados con los estereotipos acerca de los beneficios de la profesión del docente (vacaciones y sueldo; estabilidad económica o laboral); es el caso de los llamados «pragmáticos»; y, por último, si la docencia se contempla en posición final, los motivos son propios de la categoría «alternativos», porque no quieren desechar ninguna posibilidad a priori.

En nuestra opinión, el hecho de situar la enseñanza en Educación Secundaria de acuerdo con un orden de preferencia revela cómo el cambio desde la posición preferente a la intermedia y a la final supone una diferencia cualitativa sustancial en cuanto a los motivos de elección. Este posicionamiento podría ayudar a comprender la inconsistencia existente entre algunos trabajos previos, en los que los motivos a los que se alude para la elección de la docencia en Educación Secundaria como salida profesional son totalmente contradictorios. Recordemos que existen investigaciones que

aseguran que los motivos por los que los sujetos eligen la docencia como profesión son vocacionales (González Blasco y González-Anleo, 1993); otras que aseguran que la elección tuvo que ver con las dificultades del mundo laboral (Esteve, 1997; Sepúlveda y Rivas, 2000; García y Martínez, 2001), y otras en las que la elección se relaciona con la seguridad que ofrece la situación de funcionario (Berger, 1976). De nuestro trabajo se deriva que una proyección de la docencia en posición preferente se relacionará con motivos «vocacionales» y propios de «simpatizantes» (en la línea de trabajos como los de González Blasco y González-Anleo, 1993) y que, sin embargo, a medida que esta salida profesional se aleja de las preferencias del sujeto, los motivos tienen más que ver con afirmaciones «pragmáticas» y, en última instancia, como respuestas que hemos denominado «alternativas» (en la línea de investigaciones como las de Berger, 1976; Esteve, 1997; Sepúlveda y Rivas, 2000; y García y Martínez, 2001).

## Conclusiones y consideraciones finales

En este trabajo, se han analizado los motivos que acompañan a la intención del universitario de ciencias de elegir la docencia en Educación Secundaria como salida profesional futura.

El recorrido por la literatura con el que comenzábamos alerta de la necesidad de llevar a cabo un análisis de las motivaciones por las cuales una persona decide tomar el camino de la docencia en Educación Secundaria para su desarrollo profesional. El estado en el que se encuentra la enseñanza en esta etapa educativa hoy requiere de profesionales comprometidos cuya máxima motivacional y profesional sea la educación. Si ya es complicado ejercer como docente cuando estás ante un reto profesional que te motiva y te gusta, ni que decir tiene que la dificultad se duplica con la desmotivación y la frustración que supone saber que te encuentras en dicha profesión porque el mundo laboral no te ha dado otra alternativa, por lo puramente material, etc.

Desde nuestro punto de vista, este análisis de las motivaciones previas a la docencia tendría que extrapolarse al período que denominamos de formación previa. El acceso al período de FIPS debería incluir esta valoración de motivos y endurecerse de acuerdo con la consecución de un mayor compromiso por parte de los aspirantes a docentes, así como tomar como referentes aquellas experiencias en las que los resultados están siendo satisfactorios, como por ejemplo en la Universidad Autónoma de Madrid, donde

la puesta en marcha de cursos para la obtención del Título de Especialización Didáctica (TED) están logrando profundizar en la formación psicopedagógica y científica o disciplinar, partiendo de una titulación universitaria de carácter disciplinar y un diploma complementario de carácter didáctico, que tiene carácter de postgrado y que debe homologar la Formación Inicial del Profesorado de España con el resto de países de la Unión Europea, mediante la integración en el Espacio Europeo de Educación Superior.

Para el nuevo modelo de FIPS que se prevé, el modelo Master, es muy probable que el perfil del universitario que se decida a cursarlo, sea muy diferente del perfil del que decide hacer el CAP, aunque sólo sea por la ampliación temporal que entraña. No obstante, sería interesante tener en cuenta datos como los que presentamos e incluir en los requisitos de acceso una parte importante de trabajo con los candidatos en el que se analice la motivación de los mismos y se informe del nivel de compromiso que supone la educación de los futuros ciudadanos hoy.

### **Limitaciones del trabajo y perspectivas de futuro**

Este trabajo, realizado en el contexto de una investigación académico-formativa, presenta una serie de limitaciones que, por otro lado, nos permiten orientar posibles nuevos estudios en el ámbito de la FIPS.

En futuros trabajos, pretendemos utilizar una versión modificada del cuestionario aquí presentado, adaptada al alumnado que se encuentra cursando el CAPY, en última instancia, llevar a cabo entrevistas con los expertos y con profesorado de Educación Secundaria en activo acerca de la viabilidad, ventajas, desventajas, etc., del modelo de FIPS más previsible así como del papel de la formación psicopedagógica en este modelo. Somos conscientes de las limitaciones de nuestra muestra y, por tanto, tenemos que ser cautos y evitar la generalización de los datos que presentamos, por lo que el siguiente paso sería el de llevar a cabo una valoración más amplia y en diferentes contextos así como tener en cuenta en los análisis el importante factor que supone la diferencia de género.

La finalidad principal será la elaboración de claves y directrices para formular propuestas específicas de carácter psicopedagógico y su articulación de cara a un perfil ideal de formación inicial basada en competencias, contenidos y procesos. Cada vez son más los autores que hacen referencia a la necesidad de una formación psicopedagógica por parte de los docentes en la Educación Secundaria, debido a las características tan especiales de esta etapa y a los cambios socio-culturales por los que pasa la sociedad actual.

## Referencias Bibliográficas

- AMIEL, R. (1980). Équilibre mental, fatigue psychique et vocation enseignante. En VV.AA., *Équilibre ou fatigue par le travail* (77-82). París : ESF .
- BERGER, I. (1976). Psicología de los enseñantes. En VV.AA., *Tratado de ciencias pedagógicas* (147-166), vol. IX. Barcelona: Oikos Tau.
- BUENDÍA, L., COLÁS, P.Y HERNÁNDEZ, F. (1998). *Métodos de investigación en psicopedagogía*. Madrid: Mc Graw Hill.
- CAMPANARIO, J. M. (1998). ¿Quiénes son, qué piensan y saben los futuros profesores de ciencias? *Revista interuniversitaria de formación del profesorado*, 33, 121-140.
- CANO, E. (2005). *Cómo mejorar las competencias de los docentes: Guía para la auto-evaluación y el desarrollo de las competencias del profesorado*. Barcelona: Graó.
- CARDONA, M. C. (2002). *Introducción a los métodos de investigación en educación*. Madrid: EOS.
- COHEN, J. (1960). A Coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20, 37-46.
- CÓRDOBA, F., ORTEGA, R. Y MONKS, C. (2006). La formación inicial del profesorado de Secundaria en Europa: el modelo múltiple del Reino Unido. *Revista de Ciencias de la Educación*, 206, 167-176.
- CÓRDOBA, F., ORTEGA, R. Y PONTES, A. (2006). La docencia en Educación Secundaria como salida profesional para el alumnado universitario de Humanidades y Ciencias Sociales: un estudio sobre intereses y motivos. *Revista de Investigación Educativa*, 22, (aceptado para su publicación en nov, de 2006).
- CÓRDOBA, F., ORTEGA, R. Y RODRÍGUEZ, A. J. (2005). La formación inicial del profesorado de Educación Secundaria: un estudio en curso. En J.A. DEL BARRIO, M<sup>a</sup> I. FAJARDO, F. VICENTE, A. VENTURA Y I. RUIZ (eds.), *Nuevos Contextos Psicológicos y Sociales en Educación* (663-672). Barcelona: Psicoex.
- CÓRDOBA, F., ORTEGA, R. Y ROMERA, E. (2006). *La formación psicopedagógica en la formación inicial del profesorado de Secundaria: un estudio con candidatos*. En CD de Actas del I Congreso Internacional de Psicopedagogía, Melilla, 3-6 de mayo.
- CÓRDOBA, F. Y ORTEGA, R. (2006). La formación inicial del profesorado de Secundaria en nuestro país y el papel de la formación psicopedagógica: un recorrido histórico desde los 70. *ED. UCO. Revista de Investigación Educativa*, 2, 205-215.
- ESTEVE, J. M. (1997). *La formación inicial de los profesores de secundaria. Una reflexión sobre el curso de cualificación pedagógica*. Barcelona: Ariel Educación.

- GARCÍA BARROS, S. Y MARTÍNEZ LOSADA, C. (2001). Las ideas de los alumnos del CAP, punto de referencia para reflexionar sobre formación docente. *Revista Interuniversitaria de Formación del Profesorado*, 40, 97-110.
- GIMENO, J. (1988). Proyecto de reforma. *Cuadernos de Pedagogía*, 161, 47-51.
- GONZÁLEZ BLASCO, P. Y GONZÁLEZ-ANLEO, J. (1993). *El profesorado en la España actual. Informe sociológico sobre el profesorado no universitario*. Madrid: S.M.
- LEÓN, A. (1980). La profesión docente; motivaciones, actualización de conocimientos y promoción, en M. DEBESSE y G. MIALARET. *La función docente*. Barcelona: Oikos-Tau, 29-54.
- LEY ORGÁNICA 1/1990, de 3 de octubre, de Ordenación General del Sistema Educativo (LOGSE). *Boletín Oficial del Estado (España)*, 4 de octubre de 1990.
- LEY ORGÁNICA 10/2002, de 23 de diciembre, de Calidad de la Educación (LOCE). *Boletín Oficial del Estado (España)*, de 24 de diciembre de 2002.
- LEY ORGÁNICA 2/2006, de 3 de mayo, de Educación (LOE). *Boletín Oficial del Estado (España)*, de 4 de mayo de 2006.
- MEC (2006). *Propuesta de título universitario oficial de Master según R.D. 56/2005, de 21 de enero*. Consultado 4 abril de 2006, de <http://www.mec/universidades.es>
- MEIRIEU, PH. (1991). *Le choix d'éduquer*. París: ESE
- MICHAVILLA, F. (2004). *Contra la contrarreforma universitaria: crónica esperanzada de un tiempo compulsivo*. Madrid: Tecnos.
- MORENO, J. M. (2006). Profesorado de Secundaria y Calidad de la Educación: Un marco de opciones políticas para la formación y el desarrollo profesional docente. *Profesorado. Revista de currículum y formación del profesorado*, 10, 1. Consultado el 9 de febrero de 2007, de <http://www.ugr.es/~recfpro/>
- REAL DECRETO 1692/1995, 20 de octubre, por el que se regula el Título Profesional de Especialización Didáctica. *Boletín Oficial del Estado (España)*, de 9 de noviembre de 1995.
- REAL DECRETO 118/2004, de 23 de enero, por el que se regula el Título de Especialización Didáctica. *Boletín Oficial del Estado (España)*, de 4 de febrero de 2004.
- SEPÚLVEDA, M<sup>a</sup>. P. Y RIVAS, J. I. (2000). *Biografías profesionales*. Universidad de Málaga: Elementos Auxiliares de Clase 70.
- STRAUSS, A. (1987). *Qualitative analysis for social scientist*. MA: Cambridge, University Press.
- VARELA, J. Y ORTEGA, F. (1984). *El aprendizaje de maestro*. Madrid: MEC.

**Dirección de contacto:** Francisco Córdoba Alcaide. Facultad de Ciencias de la Educación. Departamento de Psicología. Avenida San Alberto Magno, s/n, 14004 Córdoba, España. E-mail: [z42coraf@uco.es](mailto:z42coraf@uco.es)

## ANEXO I: Cuestionario de investigación

En la actualidad, el hecho de que una sociedad disponga de universitarios y universitarias bien formados e informados es señal de progreso y desarrollo. Por lo tanto, el considerar sus ideas y aportaciones en la investigación sobre la formación docente puede ser muy positivo y enriquecedor. Para nuestro trabajo, en el campo de la formación inicial del profesorado de Educación Secundaria nos serían de gran utilidad las opiniones e ideas que nos puedan aportar personas como usted. El cuestionario que le presentamos a continuación es breve y no le llevará mucho tiempo completarlo. Le recordamos que es totalmente anónimo y que no existen respuestas correctas o incorrectas. Confiamos en su sinceridad y agradecemos su colaboración.

<b>¿Qué licenciatura cursa actualmente?</b>	
<b>¿En qué curso se encuentra matriculado/a? (Si se encuentra matriculado/a en asignaturas de diferentes cursos, indíquenos el curso en el que se haya matriculado de más asignaturas)</b>	
<b>Indique el género al que pertenece con una X: Masculino ( ) Femenino ( )</b>	
<b>Escriba, por orden de preferencia, un máximo de 5 salidas profesionales que le proporcione la licenciatura que está cursando, desde la que más le atraiga o interese ( N°1) hasta la que menos (N°5).</b>	<b>Preferencias (N°)</b>
	<b>1</b>
	<b>2</b>
	<b>3</b>
	<b>4</b>
	<b>5</b>
<b>¿Ha incluido la enseñanza en Educación Secundaria entre las salidas profesionales de su carrera? (Marque con una X y explique brevemente su respuesta).</b>	
<input type="checkbox"/> Sí se ha incluido. ¿Por qué?  <input type="checkbox"/> No se ha incluido. ¿Por qué?	



# Efectos de la cesión de responsabilidad de la evaluación en la motivación y la competencia percibida en el aula de educación física

## Effects of transferring responsibility for evaluation of motivation and perceived competence in physical education lessons

Juan Antonio Moreno Murcia

*Universidad de Murcia. Facultad de Ciencias del Deporte. Murcia, España.*

José Antonio Vera Lacárcel

*Universidad de Murcia. Unidad de Investigación en Educación Física y Deportes. Murcia, España*

Eduardo Cervello Gimeno

*Universidad Miguel Hernández de Elche. Facultad de Ciencias Sociales y Jurídicas de Elche. Departamento de Psicología de la Salud. Elche (Alicante), España*

### Resumen

El estudio se centró en conocer los efectos que una metodología basada en la cesión de responsabilidad en la evaluación tenía en la motivación y la competencia percibida en alumnado de sexto de primaria en el aula de educación física. Se presenta un estudio preliminar con 96 escolares de edades comprendidas entre los 11 y 12 años, divididos en dos grupos experimentales ( $n = 49$ ), y dos grupos control ( $n = 46$ ) sin experiencia en la utilización de una metodología participativa centrada en la cesión de responsabilidad al alumnado. El procedimiento a seguir al comienzo de cada unidad didáctica en el grupo experimental fue la utilización de un estilo de enseñanza basado en la instrucción directa. La cesión de responsabilidad al alumnado en la organización del aula cambió el estilo de enseñanza hacia la búsqueda y el descubrimiento. Mientras, el grupo control realizó todas las unidades didácticas a través de un

estilo de enseñanza basado únicamente en la instrucción directa. Los instrumentos administrados fueron el cuestionario de percepción de éxito (POSQ), el *cuestionario de percepción de estrategias motivacionales en las clases de educación física (CPEMEF)*, el *factor competencia percibida de la PSQ* y el *factor rol del alumno en la cesión de responsabilidad en la evaluación de la (ERAEFF)*. La intervención tuvo lugar durante un curso escolar con pretest y postest. Los resultados presentaron una disminución de la orientación disposicional al ego y el clima ego, así como la pérdida de las diferencias de género encontradas en la competencia percibida y la orientación disposicional a la tarea.

*Palabras clave:* cesión de responsabilidad, evaluación, orientación disposicional, clima motivacional, competencia percibida, educación física

### **Abstract**

The study focused on understanding the effects resulting from the application of a methodology based on the transfer of responsibility in evaluation. The effects of this methodology were analysed via the motivation and competence showed by sixth grade students in Physical Education lessons at primary school. A total of 96 students, aged between 11 and 12 years, were separated into two experimental groups (n=49) and two control groups (n=46). They had no prior experience of the use of a methodology based on the transfer of responsibility to the student. Participants' motivation and competence were evaluated via the *perception of success questionnaire (POSQ)*, the *perception of motivational strategies during the physical education lesson questionnaire (CPEMEF)*, the *perceived competence of the questionnaire (PSQ) factor*, and the *student's role in the transfer of responsibility for the evaluation on the ERAEFF scale*. The questionnaires were delivered during the school year, with pre-test and post-test. The results showed a decrease in 'ego goal orientation' and 'ego climate', as well as the loss of differences owing to gender found in the perceived competence, and the dispositional orientation in the task.

*Key Words:* transfer of responsibility, evaluation, dispositional orientation, motivational climate, perceived competence, physical education.

## Introducción

La autonomía juega un papel importante en el impacto que tiene sobre los factores situacionales que afectan a la motivación. Así, las investigaciones (Vallerand y Rousseau, 2001) han informado de niveles más altos de motivación intrínseca en aquellos alumnos a los que se les permitía elegir las tareas en el aula de Educación Física. Esta participación construida desde la autonomía de acción resulta necesaria si el aumento en la valoración de las clases se ha visto mejorado cuando el docente ha generado un clima motivacional que proporciona autonomía y responsabilidad al alumno (Cervelló, Jiménez, Del Villar, Ramos, y Santos-Rosa, 2004; Duda, 2001; Jiménez, Iglesias, Santos Rosa, y Cervelló, 2003; Roberts, 2001). De este modo, las oportunidades de los alumnos para participar en la evaluación han presentado efectos positivos en la motivación intrínseca cuando se han planteado desde estilos de enseñanza que suministran a los alumnos la oportunidad de elegir y de tomar decisiones (Hassandra, Goudas, y Chroni, 2003).

Pero, la determinación de la responsabilidad a través del compromiso del alumno resulta preocupante para el docente cuando a los alumnos se les ceden responsabilidades, no solamente por los problemas iniciales en el control de la clase sino *porque* el rumbo que la participación toma, puede generar desconcierto en el contexto de logro. La relación entre el sistema instruccional y el organizativo que el grupo de trabajo genera puede balancear las orientaciones motivacionales desde la preocupación por la maestría en la tarea a la comparación con el resto de compañeros, lo cual, afecta al sentimiento de autoeficacia.

La necesidad de percibir con eficacia las tareas realizadas generando sentimiento de competencia fue expuesta por Bandura (1987) como el motivo que más influencia tenía en la vida del alumno. Autores como Gutiérrez (2000) y Martens (1996) indican que las experiencias negativas en la práctica de actividad física hacen que las personas se consideren menos competentes. En este sentido, la teoría de las Metas de Logro (Nicholls, 1989) *sugiere* que la intervención de un alumno en un contexto de logro es evaluada de forma diferente. *Hay alumnos* que perciben que el éxito en la actividad se consigue mediante la superación de los demás y la demostración de mayor capacidad y rendimiento (*orientación hacia el ego*), existiendo una diferenciación entre la habilidad y la capacidad personal, otros perciben este éxito definido por el dominio de la tarea, el esfuerzo y el progreso personal (*orientación hacia la tarea*), considerando que no existen diferencias entre la habilidad y el esfuerzo, por lo que un gran esfuerzo supone una gran habilidad (Li y Lee, 2004).

Los estados de implicación del alumno cuando se cede responsabilidad en las tareas de evaluación parecen constituir junto con el clima de aula la motivación adecuada en el desarrollo de programas donde prima la conducta responsable en el proceso de enseñanza-aprendizaje. Así, las investigaciones establecen relaciones significativas entre la *orientación hacia la tarea* y un clima que implica a la tarea y entre la *orientación hacia el ego* y un clima que implica al ego (Digelidis, Papaioannou, Lapidis, y Christoulouidis, 2003; Escartí y Gutiérrez, 2001; Weigand y Burton, 2002). Igualmente Papaioannou, Marsh, y Theodorakis (2004) encontraron que la transmisión de *un clima que implicaba al ego disminuía la orientación hacia la tarea, y que la transmisión de un clima que implicaba a la tarea disminuía la orientación hacia el ego. La orientación hacia la tarea se encuentra relacionada con un nivel de motivación intrínseca más elevado ya que la actividad es un fin en sí misma mientras que la orientación hacia el ego es un medio para conseguir el fin* (Cervelló, Escartí, y Balagué, 1999).

Los mecanismos para responsabilizar a los alumnos en la evaluación hacen referencia al interés en la organización e implicación de las tareas de aprendizaje. Las investigaciones de Chen y Shen (2004) y Wallhead y Ntoumanis (2004), *mostraron* que el interés es más elevado cuando el aprendizaje se orienta *hacia la* tarea que cuando se hace al rendimiento. Los alumnos que presentaban un bajo nivel de habilidad y se les orientaba a la consecución de una meta perdían interés, evitaban las tareas de difícil aprendizaje, así como atribuían el éxito o el fracaso a la norma impuesta en la habilidad. Mientras si el alumno se hacía partícipe de su maestría incrementaba la motivación hacia la participación, pues definía el éxito en base al esfuerzo individual. Así, numerosas investigaciones apuntan hacia que los individuos *orientados hacia la tarea* perciben el deporte y la actividad física como una actividad que fortalece la capacidad de cooperación, la responsabilidad social y la motivación intrínseca (Duda, Chi, Newton, Walling, y Catley, 1995; Escartí y Brustad, 2000; Wang y Biddle, 2001).

Las estrategias sobre cesión de responsabilidad irán encaminadas a orientar la enseñanza al desarrollo de la habilidad en la tarea, al interés por la construcción de las actividades y a la valoración del esfuerzo personal y grupal en el desarrollo de la habilidad. Esta cesión de responsabilidad al alumno empieza por la elección de los objetivos, su discusión y puesta en práctica. Como apuntaban Del Valle, Ruiz, Vega, Velázquez y Díaz (2005) *el profesorado* debe desarrollar en sus alumnos la capacidad de reflexionar críticamente sobre sus propias acciones y su aprendizaje. Esta forma de construcción del conocimiento parece tener más efectividad si los alumnos aprenden a evaluar las tareas (Barbier, 1993).

Por otro lado, las investigaciones que relacionan el género con los criterios de percepción de éxito, revelan una mayor *orientación hacia el ego* por parte de los chicos y/o una mayor *orientación hacia la tarea* en las chicas (Carr y Weigand, 2001; Escartí, Cervelló, y Guzmán, 1996; Xiang, McBride; Guan, 2004). Sin embargo, Santos Rosa (2003) encuentra que los chicos estaban más orientados *hacia la tarea* que las chicas, y Alonso, Martínez, Moreno y Cervelló (2005) apuntaban que los chicos estaban más orientados hacia el ego y la tarea que las chicas. Del mismo modo, las investigaciones muestran una relación positiva y significativa del género masculino con el clima motivacional que implica al ego y una relación positiva y significativa del género femenino con el clima que implica a la tarea (Alonso et al., 2005; Carr y Weigand, 2001; Ntoumanis y Biddle, 1999; Santos-Rosa, 2003). El predominio en los varones a que el logro y la motivación este determinado por la comparación con el resto del grupo manifiesta que el estilo de interacción que chicos y chicas tienen con respecto a las actividades de enseñanza presenta diferencias, las chicas prefieren la interacción con sus compañeros al dominio de la habilidad, lo que afecta a la competencia percibida (Ruiz, Graupera, Rico y Mata, 2004).

La cesión de responsabilidad al alumno no está exenta de esta situación, los alumnos de ambos géneros no están dispuestos a renunciar al coste emocional que para ellos ha dibujado el contexto social, reproduciendo estos roles cuando tienen que tomar decisiones. Esta participación social basada en la competitividad produce en los niños una tendencia a emitir juicios sobre su propia habilidad que tienen por objetivo la demostración de mayor capacidad, mientras que en las niñas la carencia en el tiempo de práctica, consecuencia del énfasis en la relación, produce falta de importancia en el progreso de la habilidad, lo que puede hacerse manifiesto en las decisiones de autoevaluación. Por ello, consideramos que la utilización de estrategias docentes debería considerar escenarios en los que el clima motivacional sea favorable a ambos géneros. En este sentido, Dyson, Griffin y Hastie (2004) apuntaban a un entorno cooperativo como un medio adecuado para el desarrollo y aprendizaje de habilidades. Según los resultados de estas últimas investigaciones se hipotetiza que una cesión de responsabilidad en la evaluación que de prioridad a las preferencias cooperativas influirá en la motivación y en la relación del género con la percepción de competencia motriz.

Así pues, el objetivo que la investigación persigue es analizar los efectos de la cesión de responsabilidad en la evaluación sobre los criterios de percepción de éxito, el clima de aula y la competencia motriz percibida.

## Método

### Participantes

El alumnado participante en el estudio pertenecía al último curso de Educación Primaria ( $N = 96$ ) y se encontraba dividido en cuatro aulas. La muestra se dividió en dos grupos experimentales compuesta por 49 alumnos (14 chicos, 10 chicas y 16 chicos, 9 chicas) y dos grupos control compuestos por 47 alumnos (12 chicos, 11 chicas y 13 chicos, 11 chicas). Las edades estaban comprendidas entre los 11 y 12 años ( $M = 11$  años y dos meses,  $DT = .42$ ). Los participantes eran miembros de dos colegios públicos situados en la comarca del Mar Menor de Murcia (España), próximos uno de otro y de condición económica medio-baja. De los 96 participantes, 17 no practicaban actividad física-deportiva en horario extra-escolar, 20 lo hacían una vez por semana, 35 practicaban dos o tres veces por semana, y 24 más de tres veces por semana.

El profesorado participante tenía entre ocho y diez años de experiencia en la enseñanza de la educación física. Ambos fueron alumnos de las primeras promociones de maestros especialistas en educación física en la Universidad de Murcia.

### Instrumentos

#### Cuestionario de percepción de éxito (POSQ)

Utilizamos la versión en español (Cervelló, Escartí, y Balagué, 1999; Martínez Galindo, Alonso y Moreno, 2006) del Cuestionario de Percepción de Éxito (Roberts y Balagué, 1991; Roberts, Treasure, y Balagué, 1998) para medir las orientaciones de meta disposicionales de los escolares en las clases de educación física. El cuestionario lo componían doce ítems, seis de los cuales respondían al factor «orientación hacia la tarea» del alumno (p. ej. «En clase de educación física siento que tengo éxito al trabajar duro») y los otros seis, al factor orientación hacia el ego del alumno (p. ej. «En clase de educación física siento que tengo éxito cuando derroté a los demás»). Los ítems van precedidos de la frase «En las clases de educación física siento que tengo éxito...». Las respuestas a dicho cuestionario eran cerradas y respondían a una escala tipo Likert que oscilaba desde 0, valor que correspondía a *totalmente en desacuerdo*, a 10 valor correspondiente a *totalmente de acuerdo* con lo que se le planteaba. Se obtuvieron alphas en la toma pre y post de .79 y .88 para la subescala tarea y de .81 y .86 para la subescala ego.

## **Cuestionario de percepción de estrategias motivacionales en las clases de educación física (CPEMEF)**

Utilizamos el Cuestionario de estrategias motivacionales en las clases de educación física (CPEMEF) de Cervelló, Del Villar, Moreno y Reina (2006) para conocer la percepción de los alumnos del clima motivacional que transmite el docente. Dicho cuestionario constaba de 24 ítems, doce de los cuales hacían referencia al factor «Percepción de Estrategias Motivacionales que implicaban a la Maestría» (p. ej. «Creo que nuestro profesor confía en nuestra responsabilidad para realizar las tareas que nos propone»). El segundo factor formado por otros doce ítems, correspondía a conductas del *profesorado* que inducían a la «Percepción de Estrategias Motivacionales que implicaban al Rendimiento» (p. ej. «Mi profesor no nos deja ninguna responsabilidad»). Las respuestas a dicho cuestionario eran cerradas y respondían a una escala tipo Likert que iba de 0, valor que correspondía a *totalmente en desacuerdo*, a 10 que indicaba que el *alumnado* estaba *totalmente de acuerdo* con lo que se le planteaba. Los ítems estaban precedidos por la frase «En las clases de educación física...». La fiabilidad pre y post obtenida fue de .71 y .74 para el *clima implicante a la tarea* y .76 y .74 para el *clima implicante al ego*.

## **Competencia percibida**

Utilizamos el factor *competencia percibida* de la versión española (Moreno y Cervelló, 2005) del Physical Self-Perception Profile (Fox, 1990; Fox y Corbin, 1989). Dicho factor se componía de seis ítems (p. ej. «Soy muy bueno en casi todos los deportes y también realizando ejercicios físicos») que se respondía mediante una escala tipo Likert, cuyos rangos de puntuación oscilaban entre 0, valor que correspondía a *totalmente en desacuerdo*, y 10, valor correspondiente a *totalmente de acuerdo* con lo que se le planteaba al *alumnado*. Los ítems van precedidos de la frase «Cuando realizo actividad física...». La fiabilidad pre y post obtenida fue de .80 y .86.

## **Rol del alumnado en la cesión de responsabilidad en la evaluación**

Utilizamos el factor *rol en la cesión de responsabilidad* de la escala de responsabilidad del *alumnado* en la evaluación en educación física (Moreno, Vera, y Cervelló, 2006). El factor se componía de seis ítems (p. ej. «El maestro me pregunta la opinión que tengo de mi evaluación») que se respondía mediante una escala tipo Likert cuyos rangos de puntuación oscilaban entre 0, valor que correspondía a *totalmente en desacuerdo*, y 10, valor correspondiente a *totalmente de acuerdo*. Los ítems iban precedidos de la frase «Creo que en la clase de educación física...». La fiabilidad pre y post fue de .78 y .85.

## Procedimiento

Todo el alumnado participó voluntariamente en el estudio, quedando informados del mismo sus padres, madres o tutores. El permiso para llevar a cabo la investigación fue dado por el director del centro, los padres de los alumnos y el claustro de maestros.

Los cuestionarios administrados para analizar las variables objeto de estudio se pasaron las primeras semanas de octubre del curso 2004-05 y se volvieron a repetir las dos últimas semanas del mes de mayo, todo ello, sin la presencia en el aula del *profesorado* de educación física. Se utilizaron dos días para su cumplimentación en semanas alternas, utilizando para ello el comienzo de la clase. En la primera cumplimentación se rellenó el factor competencia percibida y el cuestionario de percepción de éxito y en la segunda el de estrategias motivacionales y el factor rol del *alumnado* en la cesión de responsabilidad. Este factor pretendía verificar que las decisiones objeto de análisis fueron cedidas por el *profesorado* y aceptadas por el alumnado. Los participantes leían el cuestionario y preguntaban todas las dudas que tuvieran, utilizando 15 minutos aproximadamente para contestar. No se encontraron dudas en la comprensión de los ítems.

El programa de intervención comenzó después de la cumplimentación de los cuestionarios. El profesorado de los grupos planificó conjuntamente siete unidades didácticas divididas en 60 sesiones que se distribuyeron a lo largo del curso escolar. Los contenidos a impartir fueron los mismos en todos los grupos. La selección del profesorado participante se realizó a partir de entrevistas llevadas a cabo antes de la investigación con el fin de seleccionar los grupos según creencias diferentes en la utilización de las estrategias de enseñanza aplicadas al aula de educación física.

En el grupo experimental se dividió cada unidad didáctica en tres momentos diferenciados: una primera fase de experiencia dirigida por el maestro basada en la reproducción del conocimiento caracterizada por un sistema de organización masivo-consecutiva donde las tareas se explicaban al conjunto de la clase y todos los alumnos las realizaban al mismo tiempo, la primera actividad, la segunda y así sucesivamente. A mitad de las sesiones que conformaban la unidad comenzaba una segunda fase donde se cedía responsabilidad a los alumnos en la elección de los objetivos de enseñanza y la construcción de las tareas, caracterizada por un sistema de organización donde la clase se dividía en varios subgrupos a partir de la similitud de objetivos elegidos. Esta elección de los objetivos se realizaba por escrito en una ficha que facilitaba el maestro para su posterior discusión. Del mismo modo, anexo a la ficha de objetivos, se administraba al alumno otra ficha donde se hacían constar las capacidades a desarrollar.

Existía un alumno-responsable que se encargaba de transmitir las impresiones del maestro sobre el trabajo del grupo.

Por último, una tercera fase donde se les pedía su colaboración en la evaluación del trabajo realizado. Para la participación en la evaluación de la unidad entregaban la ficha de objetivos, donde escribían el grado de consecución de los objetivos planteados al inicio de la unidad, y cómo consideraban su evaluación según las capacidades a observarse. Para la participación en la evaluación trimestral, se convocaba una asamblea de evaluación donde cada alumno a través de dos fichas de evaluación diseñadas para todo el curso, una individual, donde se defendía la calificación que cada alumno se daba y otra colectiva donde se vertían opiniones de los compañeros de grupo, se les cedía responsabilidad en un espacio reservado en el boletín de calificaciones.

El maestro adoptaba un papel de guía después de la primera fase, utilizando estrategias en la cesión de responsabilidad encaminadas a orientar la enseñanza al desarrollo de la habilidad en la tarea, al interés por la construcción de las actividades y a la valoración del esfuerzo personal y colectivo en el desarrollo de la habilidad. En el grupo control la elección de tareas y la evaluación eran responsabilidad del maestro.

Para comprobar si existían diferencias significativas entre los sistemas de organización sobre la gestión del tiempo, utilizamos una hoja informatizada de registro del tiempo de clase donde se analizó en una sesión por trimestre para cada uno de los estilos de enseñanza, el tiempo de atención (TAA), tiempo que el maestro dedica a explicar una tarea y corregir a los alumnos, el tiempo de organización (TOA), tiempo que se dedica a colocar y retirar material y a organizar a los alumnos en la tarea, el tiempo de actividad motriz (TAM) dedicado a practicar actividad física, y el tiempo de imprevistos (TI), pérdidas de tiempo que ocurren en las clases.

### **Diseño y análisis estadístico**

El diseño de la investigación fue de tipo cuasi-experimental pre-post con dos grupos de control ( $n = 47$ ) y dos experimentales ( $n = 49$ ). Las variables independientes que se establecieron para realizar la investigación fue la metodología empleada en el desarrollo de los contenidos de educación física (al grupo experimental se le aplicaba una metodología que comenzaba por la reproducción del conocimiento en un sistema de organización masivo-consecutivo y terminaba en la producción del conocimiento de los alumnos en un sistema de organización de varios subgrupos). Las variables dependientes fueron las orientaciones de meta, las estrategias motivacionales, la competencia motriz percibida y el papel del alumno en la cesión de responsabilidad.

Se realizaron diferentes pruebas estadísticas, cálculo del coeficiente alpha de Cronbach que nos permitió conocer la consistencia interna de los instrumentos y de cada uno de los factores resultantes, igualmente se calcularon las medias y desviaciones estándares de todos los factores. Con la prueba de correlaciones bivariadas pretendíamos conocer si las relaciones entre variables habían sufrido modificaciones tras el programa de intervención. Para valorar los efectos del programa se utilizó una prueba t para muestras relacionadas.

## Resultados

### Análisis Preliminar

Se realizó un análisis intragrupo de medidas repetidas. La prueba de esfericidad (prueba de Maunchly) para comprobar la homogeneidad de varianzas entre los distintos tiempos que formaban cada variable independiente determinó un valor no significativo ( $p > .05$ ). No existían diferencias dentro de los sistemas de organización entre la distribución del tiempo de clase. Los resultados del test de comparaciones múltiples de Bonferroni mostraron diferencias significativas, la fase de reproducción del conocimiento tenía mayor tiempo de organización ( $p < .05$ ) y de atención ( $p < .001$ ) y menor tiempo de actividad motriz ( $p < .05$ ) que la fase de producción del conocimiento de los alumnos.

Se realizó un MANOVA con los grupos experimentales y las variables edad (11 y 12 años) y género considerando como variables dependientes las orientaciones de meta disposicionales, las estrategias motivacionales, la competencia motriz percibida y el papel del alumno en la cesión de responsabilidad. Se encontraron diferencias significativas (Lambda de Wilks = .69,  $F(6,89) = 3.59, p < .01$ ) en la variable género, los chicos mostraban mayores puntuaciones en la orientación disposicional a la tarea y la competencia percibida que las chicas.

Igualmente, se realizó un ANOVA por grupos (control y experimental), con los datos pre considerando como variable dependiente el papel del alumno en la cesión de responsabilidad. No se encontraron diferencias significativas ( $F = 3.56, p > .01$ ) por lo que se parte de grupos homogéneos que no muestran diferencias entre ellos en el papel que juegan los alumnos en la cesión de responsabilidad en la evaluación.

## Análisis de correlación

En la Tabla I se observa como en las diferencias de las correlaciones entre los datos pre y post destaca que la orientación hacia la tarea (pre) correlacionaba positivamente con la orientación hacia el ego y la competencia percibida, mientras que en los datos (post) lo hacía con el clima implicado hacia la tarea. Destacar también que el clima que implica hacia la tarea (pre) correlacionaba negativamente con el clima que implicaba hacia el ego y en los datos (post) lo hacía positivamente hacia la competencia percibida.

TABLA I. Coeficientes Alpha de Cronbach y correlaciones bivariadas de las variables (pre y post)

	.	1	2	3	4	5	6	7	8	9	10	11	12
1. Orientación Tarea (pre)	.79	-	.38**	.14	.14	-.01	.11	.12	.16	.32*	.20	.04	.16
2. Orientación Ego (pre)	.81	-	-	.00	.52**	-.00	.44**	.03	.34*	.25	.07	-.06	.11
3. Orientación Tarea (post)	.88	-	-	-	-.12	.24	-.06	.60**	-.29*	.22	.26	-.08	.24
4. Orientación Ego (post)	.86	-	-	-	-	-.06	.35*	-.12	.40**	.53**	.04	.02	-.07
5. Clima Tarea (pre)	.71	-	-	-	-	-	-.38**	.16	-.00	-.04	-.06	.34*	.05
6. Clima Ego (pre)	.76	-	-	-	-	-	-	-.10	.41**	.27	.08	-.22	.14
7. Clima Tarea (post)	.74	-	-	-	-	-	-	-	-.45**	.01	.47**	.08	.35*
8. Clima Ego (post)	.74	-	-	-	-	-	-	-	-	.25	.09	.00	-.15
9. Competencia (pre)	.80	-	-	-	-	-	-	-	-	-	.21	-.12	-.05
10. Competencia (post)	.83	-	-	-	-	-	-	-	-	-	-	.13	.22
11. Rol del alumno (pre)	.78	-	-	-	-	-	-	-	-	-	-	-	.24
12. Rol del alumno (post)	.85	-	-	-	-	-	-	-	-	-	-	-	-

\* $p < .01$ ; \*\* $p < .001$

## Efectos de la intervención en el programa

La Tabla II muestra los resultados para cada una de las variables dependientes antes y después de la intervención en ambos programas. La prueba t para muestras relacionadas reveló que el grupo experimental descendió significativamente en la *orientación hacia el ego* ( $p < .001$ ) y el clima implicante al ego ( $p < .001$ ), incrementando significativamente en la *orientación hacia la tarea* ( $p < .01$ ), el *clima implicante a la tarea* ( $p < .01$ ) y el rol del alumno en la cesión de responsabilidad ( $p < .001$ ). En contraste, el grupo control mostró un descenso significativo en el *clima implicante a la tarea* ( $p < .01$ ) y el rol del alumno en la cesión de responsabilidad ( $p < .001$ ).

Para comprobar los efectos de interacción en los grupos experimentales entre la edad, el género y las variables dependientes competencia percibida y orientación

disposicional a la tarea se realizó un análisis multivariante, no observando diferencias significativas ( $\Lambda$  de Wilks = .84,  $F(6,89) = 2.54, p > .01$ ). Se realizó un ANOVA por grupos (control y experimental), con los datos post considerando como variable dependiente el papel del alumno en la cesión de responsabilidad, se encontraron diferencias significativas ( $F = 8.48, p < .001$ ). Los grupos experimentales percibían la cesión de responsabilidad del educador con más énfasis ( $M = 7.01, DT = 1.57$ ) que los grupos de control ( $M = 2.45, DT = 1.66$ ).

TABLA II. Prueba t para muestras relacionadas ( $M \pm DT$  de las variables antes y después de la intervención).

Variable dependiente		Grupo experimental (n = 49)		Grupo Control (n = 47)	
		M	DT	M	DT
Orientación tarea	Pre	6.93	1.60	7.51	1.79
	Post	8.19*	1.58	7.22	7.99
Orientación ego	Pre	5.01	2.88	5.86	2.42
	Post	3.57**	1.80	6.08	2.23
Clima tarea	Pre	6.60	1.38	7.29	1.30
	Post	7.37*	1.43	6.85*	1.67
Clima ego	Pre	3.33	1.80	3.13	1.58
	Post	2.04**	1.28	3.05	1.80
Competencia percibida	Pre	6.40	2.01	6.20	2.02
	Post	6.60	1.38	6.08	2.23
Rol del alumno en la cesión de responsabilidad	Pre	4.35	2.28	3.52	1.97
	Post	7.01**	1.57	2.45*	1.66

\* $p < .01$ ; \*\* $p < .001$

## Discusión

El propósito del estudio fue analizar los efectos que la cesión de responsabilidad en la evaluación tenía sobre los criterios de percepción de éxito, el clima de aula y la competencia motriz percibida. Tras la aplicación del programa de cesión de responsabilidad se plantea la existencia de una relación positiva y significativa entre la orientación hacia la tarea y el clima implicante a la tarea y la orientación hacia el ego y el clima implicante al ego que se presenta acorde con los resultados encontrados por diferentes autores (Escartí y Gutiérrez, 2001; Weigand y Burton, 2002; Xiang y Lee, 2002; Digelidis et al., 2003; Gano-Overway y Ewing, 2004). Por otro lado, la relación negativa y significativa de la orientación hacia la tarea (post) y el clima implicante al ego (post)

y la relación negativa de la orientación hacia el ego con el clima implicante a la tarea, refleja que la transmisión de un clima implicante a la tarea puede disminuir la orientación hacia el ego, y que la transmisión de un clima implicante al ego tiene los mismos efectos con la orientación hacia la tarea (Escartí y Gutiérrez, 2001; Digelidis et al.; 2003; Papaioannou et al., 2004). Esto nos hace pensar que la utilización de un clima opuesto a las orientaciones de meta que confluyen en el aula puede afectar a un cambio en las mismas.

Así, parece que los alumnos a los que se les cede responsabilidad en la participación de las actividades de aula perciben un clima orientado a la tarea (Wallhead y Ntoumanis, 2004). En esta misma línea, la investigación de Yonemura, Fukugusakio, Yoshinaga y Takahashi (2003) apuntaba que la cesión de responsabilidad en la evaluación de la efectividad de las tareas orientaba a los alumnos al aprendizaje de la tarea.

La utilización de estrategias motivacionales centradas en la construcción de las tareas, la progresiva cesión de responsabilidad, y la autonomía en la evaluación parecen provocar en la percepción de competencia una relación cambiante en la percepción del clima de aula y las orientaciones de meta disposicionales. Los alumnos manifiestan positiva y significativamente su sentimiento de autoeficacia cuando se orientan hacia la tarea y al ego (inicial), al ego (final) y cuando perciben positiva y significativamente un clima implicante a la tarea (final) y negativamente un clima implicante al ego (final). No parece que la percepción de competencia venga definida exclusivamente por una sola orientación motivacional. Como expusieron Treasure, Duda, Hall, Roberts, Ames y Maehr (2001) los criterios de éxito pueden variar de una orientación a otra en función de la disposición del alumno, aspecto este que puede guardar relación con la forma de interactuar las variables motivaciones de aula, las orientaciones de los iguales, las familiares y las creencias personales.

Los cambios producidos tras la aplicación del programa afectan significativamente a la relación del género con la percepción de competencia y la orientación hacia la tarea. Mientras inicialmente se habían presentado diferencias significativas en la percepción de competencia, los chicos se percibían más competentes que las chicas, la cesión de responsabilidad había disminuido la percepción de autoeficacia en los chicos, y en el caso de las chicas se produce un ligero aumento no significativo que hace desaparecer las diferencias encontradas inicialmente. Referente a la orientación hacia la tarea las diferencias encontradas a favor de los chicos desaparecerían después de la cesión, continuando los chicos orientados hacia la tarea y al ego, mientras que las chicas aumentaban la orientación hacia la tarea, disminuyendo la orientación hacia el ego.

Parece que la cesión de responsabilidad produce en las chicas un aumento de la orientación motivacional a la tarea que se relaciona con una disminución de la orientación motivacional al ego, el clima implicante al ego, y al aumento de la percepción de competencia, mientras que en los chicos esta misma situación no produce aumentos significativos en la percepción de competencia. En este sentido, el estudio de Solmon, Lee, Belcher, Harrison y Wells (2003) *muestra* que cuando las chicas perciben la actividad más propia de hombres que de mujeres suelen mostrarse menos competentes. Cuando las chicas interpretan que la información que sugiere la actividad realizada es propia de chicos es probable que eviten el compromiso con dicha actividad. Por ello, la orientación hacia la tarea y las estrategias motivacionales que implican a la maestría y que configuran el proceso y el resultado de la evaluación pueden ser determinantes para la ausencia en la percepción de competencia deportiva tanto de chicas como chicos, por lo que deben ser hábilmente utilizadas. Además, el agrupamiento heterogéneo donde chicos y chicas tiene que unir intereses en el planteamiento de objetivos comunes bajo un clima de aula que implica a la tarea, puede plantear en los chicos problemas que afecten a su motivación. Esta falta de motivación se relaciona con el enfrentamiento de dos autoconceptos (Ruiz et al., 2004). Los chicos se sienten más motivados hacia actividades competitivas y las chicas hacia las cooperativas, por lo que los estilos de interacción preferidos son diferentes.

Parece que la motivación es un factor que nos ayuda a comprender las diferentes percepciones individuales que tienen lugar en los alumnos. De este modo, Li, Lee, y Solmon (2005) indican que aquellos alumnos que están más motivados se perciben más competentes. La implementación del trabajo cooperativo basado en la cesión de responsabilidad parece mejorar la motivación más en chicas que en chicos.

Finalmente, la tradición de una menor percepción de competencia deportiva en las chicas (Mañano, Ninot, y Bilard, 2004; Asçi, Asçi, y Zorba, 2005; Welk y Eklund, 2005), puede verse mejorada. El género es una subcultura que se engloba en una cultura dominante, y donde la primacía oculta de los valores masculinos hacen que la participación se haya propuesto desde este dominio, lo cual puede estar afectando a las percepciones del género femenino. Futuros trabajos deberían analizar si la diferente acogida de los géneros en la cesión de responsabilidad guarda relación con los hábitos sociales aprendidos y la edad en la que comienzan a manifestarse.

## Referencias bibliográficas

- ALONSO, N., MARTÍNEZ, C., MORENO, J. A. Y CERVELLÓ, E. (2005). *Relación del género del alumno y el tipo de centro con la motivación, disciplina, trato de igualdad y estado de flow en educación física*. V Congreso Internacional de Educación Física e Interculturalidad, Murcia, España.
- ASÇI, F. H., ASÇI, A. Y ZORBA, E. (1999). Cross-cultural validity and reliability of Physical Self-Perception Profile. *International Journal of Sport Psychology*, 30, 399-406.
- BANDURA, A. (1987). *Pensamiento y acción*. Barcelona: Martínez Roca.
- BARBIER, J. M. (1993). La evaluación de acciones. Objeto y función de la evaluación. En BARBIER (ed.), *La evaluación en los procesos de formación* (p. 260). Barcelona: Paidós.
- CARR, S. Y WEIGAND, D. A. (2001). Parental, peer, teacher, and sporting hero influence on the goal orientations of children in physical education. *European Physical Education Review*, 7, 305-328.
- CERVELLÓ, E., ESCARTÍ, A. Y BALAGUÉ, G. (1999). Relaciones entre la orientación de meta disposicional y la satisfacción con los resultados deportivos, las creencias sobre las causas de éxito en deporte y la diversión con la práctica deportiva. *Revista de Psicología del Deporte*, 8(1), 7-19.
- CERVELLÓ, E. M., JIMÉNEZ, R., DEL VILLAR, F., RAMOS, L. A. Y SANTOS-ROSA, F. J. (2004). Goals orientations, motivational climate, equality, and discipline in spanish physical education students. *Perceptual and Motor Skills*, 99, 271-283.
- CERVELLÓ, E. M., MORENO, J. A., DEL VILLAR, F. Y REINA, R. (2006). Desarrollo y validación de un instrumento de medida de las estrategias motivacionales empleadas en las clases de Educación Física. Manuscrito en revisión.
- CHEN, A. Y SHEN, B. (2004). A web of achieving in physical education: goals, interest, outside-school activity and learning. *Learning and Individual Differences*, 14, 169-182.
- DEL VALLE, S., RUIZ, L. M., VELÁZQUEZ, R. Y DÍAZ, P. (2005). Sabemos que contenidos aprenden los escolares en Educación Física? *European Journal of Human Movement*, 13, 87-106.
- DIGELIDIS, N., PAPAIOANNOU, A., LAPARIDIS, K. Y CHRISTODOULIDIS, T. (2003). A one-year intervention in 7th grade physical education classes aiming to change motivational climate and attitudes towards exercise. *Psychology of Sport and Exercise*, 4, 195-210.
- DUDA, J. L., CHI, L., NEWTON, M., WALLING, M. D. Y CATLEY, D. (1995). Task and ego orientation and intrinsic motivation in sport. *International Journal of Sport Psychology*, 26, 40-63.

- DUDA, J. (2001). Achievement goal research in sport: pushing the boundaries a clarifying some misunderstandings. En G. ROBERTS (ed.), *Advances in motivation in sport and exercise* (129-182). Champaign, IL: Human Kinetics.
- DYSON, B., GRIFFIN, L. L. Y HASTIE, P. (2004). Sport education, tactical games, and cooperative learning: Theoretical and pedagogical considerations. *Quest*, 56, 226-240.
- ESCARTÍ, A., CERVELLÓ, E. Y GUZMÁN, J. F. (1996). La orientación de metas de adolescentes deportistas de competición y la percepción de los criterios de éxito deportivo de los otros significativos. *Revista de Psicología Social Aplicada*, 6(2), 27-42.
- ESCARTÍ, A. Y BRUSTAD, R. (2000). *El estudio de la motivación deportiva desde la perspectiva de la teoría de metas*. I Congreso Hispano-Portugués de Psicología. Santiago de Compostela, España.
- ESCARTÍ, A. Y GUTIÉRREZ, M. (2001). Influence of the motivational climate in physical education on the intention to practice physical activity or sports. *European Journal of Sport Science*, 1(4), 1-12.
- FOX, K. R. Y CORBIN, D. (1989). The Physical Self-Perception Profile: Development and preliminary validation. *Journal of Sport and Exercise Psychology*, 11, 408-430.
- GANO-OVERWAY, L. A. Y EWING, M. E. (2004). A longitudinal perspective of the relationship between perceived motivational climate, goal orientations, and strategy use. *Research Quarterly for Exercise and Sport*, 75, 315-325.
- GUTIÉRREZ, M. (2000). Actividad física, estilos de vida y calidad de vida. *Revista de Educación*, 77, 5-14.
- HASSANDRA, M., GOUDAS, M. Y CHRONI, S. (2003). Examining factors associated with intrinsic motivation in physical education: a qualitative approach. *Psychology of Sport and Exercise*, 4, 211-223.
- JIMÉNEZ, R., IGLESIAS, D., SANTOS ROSA, F. J. Y CERVELLO, E. M. (2003). *Análisis de la relación entre el clima motivacional, las orientaciones de meta y la igualdad de trato en función del género en las clases de Educación Física*. IV Congreso Internacional sobre la enseñanza de la Educación Física y el Deporte Escolar. Santander, España.
- LI, W. Y LEE, A. (2004). A review of conceptions of ability and related motivational constructs in achievement motivation. *Quest*, 56, 439-461.
- LI, W. Y LEE, A. M. Y SOLMON, M. A. (2005). Relationship among dispositional ability conceptions, intrinsic motivation, perceived competence, experience, persistence and performance. *Journal of Teaching in Physical Education*, 24, 51-65.
- MAÍANO, C., NINOT, G. Y BILARD, J. (2004). Age and gender effects on global self-esteem and physical self-perception in adolescents. *European Physical Education Review*, 10, 53-69.

- MARTENS, R. (1996). Turning kids on to physical activity life time. *Quest*, 48(3), 303-310.
- MARTÍNEZ GALINDO, C., ALONSO, N. Y MORENO, J.A. (2006). *Análisis factorial confirmatorio del Cuestionario de Percepción de Éxito (POSQ)*. IV Congreso de la Asociación Española de Ciencias del Deporte. A Coruña, Xunta de Galicia.
- MORENO, J. A. Y CERVELLÓ, E. (2005). Physical Self-Perception in Spanish Adolescents: Gender and Involvement in Physical Activity Effects. *Journal of Human Movement Studies*, 48, 291-311.
- MORENO, J.A., VERA, J.A. Y CERVELLÓ, E. (2006). Evaluación participativa y responsabilidad en Educación Física. *Revista de Educación*, 340, 731-754.
- NICHOLLS, J. G. (1989). *The competitive ethos and democratic education*. Cambridge, MASS: Harvard University Press.
- NTOUMANIS, N. Y BIDDLE, S. (1999). A review of motivational climate in physical activity. *Journal of Sport Sciences*, 17, 643-665.
- PAPAIANOANNOU, A., MARSCH, H. W. Y THEODORAKIS, Y. (2004). A multilevel approach to motivational climate in physical education and sport settings: and individual or a group level construct. *Journal of Sport and Exercise Psychology*, 26, 90-118.
- ROBERTS, G. C. Y BALAGUE, G. (1991). *The development and validation of the Perception of Success Questionnaire*. Paper presented at the FEPSAC Congress, Cologne, Germany.
- ROBERTS, G. C., TREASURE, D. C. Y BALAGUÉ, G. (1998). Achievement goals in sport: The development and validation of the Perception of Success Questionnaire. *Journal of Sports Sciences*, 16, 337-347.
- ROBERTS, G. C. (2001). Understanding the dynamics of motivation in physical activity: the influence of achievements goals on motivational process. En G. C. ROBERTS (ed.), *Advances in motivation in sport and exercise* (pp. 1-50). Champaign, IL: Human Kinetics.
- RUIZ, L. M., GRAKUPERA, J. L., RICO, I. Y MATA, E. (2004). Preferencias participativas en Educación Física de los chicos y chicas de la Educación Secundaria mediante la escala GR de participación social en el aprendizaje. *European Journal of Human Movement*, 12, 151-168.
- SANTOS-ROSA, F.J. (2003). *Motivación, ansiedad y flow en jóvenes tenistas*. Tesis doctoral, Universidad de Extremadura. Facultad de Ciencias del Deporte, Cáceres.
- SOLMON, M. A., LEE, A. M., BELCHER, D., HARRISON, L. Y WELLS, L. (2003). Beliefs about gender appropriateness, ability, and competence in physical activity. *Journal of Teaching in Physical Education*, 22, 261-279.

- TREASURE, D. C., DUDA, J. L., HALL, H. K., ROBERTS, J. C., AMES, C. Y MAEHR, M. L. (2001). Clarifying misconceptions and misrepresentations in achievement goal research in sport. A response to Harwood, Hardy, and Swain. *Journal of Sport and Exercise Psychology*, 23, 317-329.
- VELÁZQUEZ BUENDÍA, R. Y HERNÁNDEZ ÁLVAREZ, J. L. (2004). Evaluación en educación y evaluación del aprendizaje en educación física. En J. L. HERNÁNDEZ ÁLVAREZ, R. VELÁZQUEZ BUENDÍA, D. ALONSO CUIEL, F. J. CASTEJÓN OLIVA, I. GAROZ PUERTA, C. LÓPEZ CRESPO, A. LÓPEZ RODRÍGUEZ, A. MALDONADO RICO Y M. R. MARTÍNEZ GORROÑO (eds.), *La evaluación en Educación Física: Investigación y práctica en el ámbito escolar* (pp. 11-33). Barcelona: Grao.
- VALLERAND, R. J. Y ROSSEAU, F. L. (2001). Intrinsic motivation in sport and exercise. En R. N. SINGER, H. A. HAUSENBLAS Y C. M. JANELLE (eds.), *Handbook of sport psychology* (394-398). New York: John Wiley & Sons.
- WALLHEAD, T. T. Y NTOUMANIS, N. (2004). Effects of sport education intervention on student' motivational responses in physical education. *Journal of Teaching in Physical Education*, 23, 4-18.
- WANG, C. K. J. Y BIDDLE, S. J. H. (2001). Young people's motivational profiles in physical activity: a cluster analysis. *Journal of Sport and Exercise Psychology*, 23, 1-22.
- WEIGAND, D. A. Y BURTON, S. (2002). Manipulating achievement motivation in Physical Education by manipulating the motivational climate. *European Journal of Sport Science*, 2(1), 1-14.
- WELK, G. J. Y EUKLUND, B. (2005). Validation of the children and youth physical self perceptions profile for young children, *Psychology of Sport and Exercise*, 6, 51-65.
- XIANG, P. Y LEE, A. (2002). Achievement goals, perceived motivational climate, and students' self-reported mastery behaviours. *Research Quarterly for Exercise and Sport*, 73, 58-65.
- XIANG, P., MCBRIDE, R. Y GUAN, J. (2004). Children's motivation in elementary physical education: A longitudinal study. *Research Quarterly for Exercise and Sport*, 75, 71-80.
- YONEMURA, K., FUKUGUSAKIO, Y., YOSHINAGA, T. Y TAKAHASHI, T. (2003). Effects of Momentum and climate in Physical Education class on students' formative evaluation. *International Journal of Sport and Health Science*, 2, 25-33.

**Dirección de contacto:** Juan Antonio Moreno Murcia. Universidad de Murcia. Facultad de Ciencias del Deporte. Parque Almansa. 30730 San Javier (Murcia), España. E-mail: [morenomu@um.es](mailto:morenomu@um.es)



## **Ensayos e informes**



# Los nuevos retos del aprendizaje social en niños con necesidades educativas especiales. El aprendizaje en comunidad o la comunidad de aprendizaje en el aula

## The new challenges of social learning for children with special educational needs. Learning in the community, or the community of learning in the classroom

Rafael Pérez Galán

*Universidad de Málaga. Departamento de Didáctica y Organización Escolar. Málaga, España*

### Resumen

Este artículo presenta y apuesta por un aprendizaje basado en la cooperación y donación de significados entre alumnos en procesos de intercambio e interacción social permanente. El aula como espacio vital y psicológico donde se sitúan alumnos con experiencias, conocimientos y expectativas diferentes no puede ser un espacio único y cerrado a las demandas y exigencias sociales; debe ser un subsistema que favorezca y promueva procesos de construcción personal en cooperación con otras instancias, y que tenga en cuenta intereses y exigencias distintas según las conciencias individuales con las que comparte tiempo, espacio y realidades cercanas. Para ello, presentamos la comunidad de aprendizaje, aprendizaje de colaboración, el proyecto de trabajo, aprendizaje dialógico, aprendizaje cooperativo, como instrumentos básicos con los que poder reinterpretar los nuevos retos educativos de esta sociedad del conocimiento y de la información, con el propósito de redescubrir un conocimiento que no excluya socialmente a nadie, considere el principio de normalización y de necesidad educativa especial y aporte significado, interés y funcionalidad ante los nuevos desafíos del siglo XXI. Un aprendizaje y conocimiento con estas características sólo puede cimentarse en estrategias diversas y almas heterogéneas que conlleven procesos de negociación y reinterpretación de significados en el aula, donde tengan cabida, siempre, las interpretaciones y aportaciones de

los otros, los sentimientos, afectos de todos en contextos de diversificación y de apuesta solidaria por lo relevante de cada uno, sin discriminación alguna. Donde cada uno observe la praxis de la realidad desde su particular forma de interpretar el mundo.

*Palabras clave:* comunidad de aprendizaje, aprendizaje «dialógico», escuela comprensiva, proyecto de trabajo y escuela inclusiva.

### **Abstract**

This article presents and defends a learning experience based on the cooperation and negotiation of meanings among students in constant processes of exchange and social interaction. The classroom is understood as a vital and psychological space where a group of individuals with different experiences, background knowledge and expectations live together and which cannot be considered an isolated closed space opposed to our social demands and requirements; it must be a subsystem which favours and promotes processes of personal growing in cooperation with other organizations. It must also consider individual interests and demands from its members with whom it shares time, space and a common reality. In order to achieve this, basic instruments are proposed, such as the community of learning, collaborative learning, projects, communicative learning and cooperative learning. The aim is to reinterpret the new educational challenges of the current society of knowledge and information with a view to rediscovering a new knowledge with no social exclusion, which considers the principle of normalization and special educational needs, and brings forward meaning, interest and practice, facing the new challenges of the twenty-first century. Learning and knowledge with these characteristics can only be consolidated via varied strategies and heterogeneity, which entail processes of negotiation and reinterpretation of meaning in class, where others' interpretations and contributions, as well as feelings and affection towards each other, are always accepted in contexts of diversification and empathetic support of relevant aspects of the individuals with no discrimination at all. It is about constructing a setting where everybody observes the practice of reality from their particular way of interpreting the world.

*Key Words:* community of learning, communicative learning, comprehensive school, projects, inclusive school.

## Introducción

En la sociedad del conocimiento y de la información hay cosas que han perdido vigencia y valor. Si nos situamos en la escuela, no tiene sentido que el alumno espere que alguien le cuente cosas, que se conforme con los referentes culturales que le vienen dados por unos o por otros, que adopte una postura pasiva ante el manejo de la información, o que el aula sea el lugar insoslayable de aprendizaje. Todo ha cambiado. Lo importante radica en que la cultura se adquiere en diferentes contextos y fuera del aula, que el profesor debe ser un provocador de situaciones, y que lo relevante está fuera del contexto escolar y no dentro. Por ello nos atrevemos a decir que cada vez más el aprendizaje depende menos de lo que ocurre en el aula y cada vez más de la correlación entre lo que ocurre en el aula, el domicilio y la calle.

A partir de estas consideraciones, presentamos la comunidad de aprendizaje como una forma distinta de situarse ante el conocimiento y como una forma diferente de organizar la escuela. El conocimiento social y la dimensión que supone aprender con los otros, adquieren una importancia tan crucial que solo puede darse si toda la comunidad educativa asume la verdadera función de educar en igualdad y justicia social.

Este tipo de enseñanza, basada en la asunción de estrategias y responsabilidades de grupo, demanda otro tipo de herramientas y formas de trabajar en el aula, encaminadas, lógicamente, a la realización de tareas conjuntas y propuestas de cooperación y colaboración entre los que más pueden y los menos capaces. Se trataría de tener en cuenta las diferencias e incluirlos a todos en los procesos de construcción de significados dentro de una misma realidad contextual. La escuela debe asumir, desde esta nueva consideración, la verdadera función social que le corresponde sin discriminaciones por cuestiones diversas: raza, género, lengua, discapacidad, etc.

## Introducción al concepto de comunidad de aprendizaje o aprendizaje en comunidad

La aparición de las comunidades de aprendizaje han respondido a lo que siempre ha sido el común denominador de todas las tendencias y estructuras teórico-prácticas en educación: al fracaso escolar y la desigualdad social.

Aunque se sabe que esta idea surge fuera de nuestro entorno, en concreto en EEUU, en España comienza a hablarse de estas propuestas educativas a finales del siglo XX, situando el verdadero principio de esta nueva escuela en que el aprendizaje debe estar al alcance de todos los niños, y no de unos pocos. Este es el verdadero punto de partida para transformar una escuela en una comunidad de aprendizaje. Por tanto, transformar una escuela en comunidad de aprendizaje significa comenzar un proyecto de transformación social y cultural de un centro educativo y de su entorno (Marín y Soler, 2000).

Esta transformación debe comenzar en la escuela dando oportunidades a todos, y que todos los alumnos tengan las mismas opciones, sin que nadie se sienta excluido.

Lo que propone la comunidad de aprendizaje frente a la adaptación de los alumnos, tanto en la estructura formal de la escuela como en la del aula, es la aceleración, entendida como una forma de compensar a la persona que parte con menos posibilidades en la aventura del saber; y, también, en otro tipo de aventuras, no menos importantes: aventura de emocionarse, de vivir, de relacionarse, de compartir, de sentir, etc.

Otra sería la propuesta de grupos interactivos en la resolución de tareas, potenciando de esta forma la perspectiva «Vigotskiana» del conocimiento, que no es otra cosa que incidir en que el conocimiento se aprende con los demás, y no sin ellos. Estos grupos deben caracterizarse por dos cosas fundamentalmente: la utilización de un aprendizaje dialógico y ser heterogéneos. Con esta dinámica se pretende fomentar las interacciones entre alumnos con el fin de conseguir que se acelere la solidaridad y desaparezca la conflictividad en el aula. El objetivo es conseguir una sociedad de la información basada en el aprendizaje dialógico mediante la participación de la comunidad en todos sus espacios, incluida el aula (Elboj, Puigdemívol, Soler y Valls, 2002). Este aprendizaje dialógico, a nivel general, supone la acción conjunta del alumnado, familiares, comunidad y profesionales de la educación en situación interactiva en la construcción humana del conocimiento. Por este motivo, ya el primer paso, el sueño del nuevo tipo de escuela es producto del diálogo y del consenso entre el mayor número de sectores implicados: profesionales del sistema escolar, profesionales de otras educaciones (educadoras y educadores sociales, centros de tiempo libre), asociaciones, familiares, empresas, alumnado y ayuntamientos. Las escuelas deben constituirse en comunidades de aprendizaje donde los maestros y adultos de la comunidad sean ellos mismos aprendices ejemplares así como habilidosos facilitadores del arte de aprender (Wells y Claxton, 2002).

La cualidad clave de esta comunidad de aprendizaje, inspirada en ese aprendizaje dialógico, será la de implicarse activa e inteligentemente en situaciones complejas, tanto personales como sociales. La escuela cobra una nueva dimensión social y cultural cuando es capaz no solo de que los alumnos sean activos partícipes en la construcción

del conocimiento sino de que también tomen partido del conocimiento valioso que se debe construir en el aula, con ayuda de personas externas y ajenas al proceso de enseñanza-aprendizaje. Desde la perspectiva educativa el sujeto humano no es un individuo en construcción, sino un sujeto que se construye (Merieu, 1998).

Todas las experiencias educativas, a nivel mundial, que están logrando éxitos en la superación de las desigualdades se basan en las características de este tipo de aprendizaje dialógico, que supone la acción conjunta del alumnado, familias, comunidad y profesionales de la educación. Su importancia actual aumenta en una sociedad de la información en la que el aprendizaje depende, principalmente y cada vez más, de todas las interacciones del alumnado y no solo de las que recibe en el aula ni solo de sus conocimientos previos (Flecha, 1997).

Estamos convencidos que con esta experiencia, con la utilización de este tipo de aprendizaje y esta nueva forma de concebir el aula, como espacio compartido por todos y espacio total de convivencia, lo que se pretende es que niños y niñas que proceden de familias con un bajo nivel cultural puedan acceder a las mismas fuentes de información y puedan conocer los mecanismos oportunos de selección y uso de la información, antes en manos de unos pocos.

En la nueva organización social, el conocimiento y la información desempeñan un papel que antes no habían desempeñado, reemplazando en algunos casos a los recursos naturales, a la fuerza o al dinero como factores de la generación y la distribución del poder (Tedesco, 2000). Este tipo de sociedades que utilizan las nuevas tecnologías de producción basadas en la utilización intensiva del conocimiento generan fenómenos de mayor igualdad, pero también de mayor desigualdad; de mayor homogeneización, pero también de mayor diferenciación. Lo que significa que las desigualdades en nuestro mundo realmente están aumentando y, lo que es más, las exclusiones, incluso con la atenta mirada de la escuela. La única forma de acabar con estas situaciones de desventaja e injusticia social sería la de concebir la escuela como un lugar y encuentro de culturas, donde se dé una transformación social, cultural y académica, empujando porque el maestro genere un nuevo pensamiento y cambio de actitud en una sociedad y escuela que necesitan de otro discurso diferente y de otra forma de actuar en el aula. Se trataría de educar a nuestros jóvenes alumnos y alumnas en nuevas formas de concebir el conocimiento y de utilizar sus potencialidades en esta era de la información. Educarlos en nuevas responsabilidades como ciudadanos del planeta, en actitudes críticas y de intervención social, en nuevas formas de solidaridad y de cooperación en un mundo profundamente desigual. Se trataría, en definitiva, de concebir el aula como comunidad de aprendizaje y el aprendizaje, siempre, en comunidad.

## Elementos esenciales en una comunidad de aprendizaje

Si queremos que la escuela y, por ende, el aula se conviertan en espacios de vivencia cultural, en lugares donde los niños puedan adquirir e intercambiar significados que les sean interesantes y útiles para ellos, debe darse una transformación radical tanto en la estructura formal como profunda de los centros educativos. De nada nos vale hacer cambios significativos en todo el entramado del currículo (cambiar objetivos, modificar contenidos, proponer metodología diversa, etc.), si verdaderamente no estamos convencidos los maestros, de que hace falta un cambio en nuestra manera de enseñar y en la forma de aprender de nuestros alumnos.

Esta nueva escuela requiere de unos elementos esenciales para que se pueda crear una convivencia en el aula que persiga fines comunes y a la vez objetivos individuales, tanto en lo meramente social como en lo personal. En primer lugar, habría que hablar de *aprendizaje cooperativo* entre alumnos, como una estrategia que va mucho más allá del hecho de agrupar a los alumnos, ya que este tipo de aprendizaje requiere una planificación detallada y compleja que, por supuesto, nada tiene que ver con sentar juntos a alumnos y alumnas para que resuelvan una tarea. El aprendizaje cooperativo, como nueva estrategia para que nuestros alumnos aprendan de forma relevante y significativa, supone un enfoque natural de la enseñanza de éstos (Pérez Galán, 2003).

Señala la preocupación por el medio social que les rodea y su dependencia del grupo de compañeros. Crea un contexto para el aprendizaje en el que los estudiantes exploran nuevas ideas, examinan sus propias posiciones y desafían sus creencias previas al contrastarlas con las experiencias de otras personas. La idea fundamental del aprendizaje cooperativo, según Slaving (1985), se basa en la propuesta de una serie de estrategias de instrucción sistemáticas y estructuradas que contribuyan a que cada grupo de alumnos, aun siendo de diferente nivel de rendimiento, de sexo opuesto, de diferentes grupos sociales, o con necesidades educativas especiales, puedan trabajar juntos en la consecución de un mismo aprendizaje. De entre las condiciones que debe reunir la actividad grupal para que pueda ser considerada como cooperativa, y que pueden encontrarse en Johnson, Johnson y Holubec (1999), destaca la interdependencia positiva; donde se hace referencia a que el éxito individual está ligado al del equipo y viceversa. El alumno depende del equipo y el equipo del alumno. Esta es la máxima de un aprendizaje que se apoya en la corresponsabilidad entre los niños dentro de un espacio común, y el desarrollo de valores tan cruciales como la solidaridad, la cooperación y la donación.

Los métodos de aprendizaje cooperativo (Monereo y Durán, 2002) son diseños didácticos que nos ayudan a convertir la actividad grupal en trabajo cooperativo. Utilizados de forma creativa y ajustados a las condiciones y necesidades de cada docente pueden facilitarnos la creación de situaciones que se benefician del poder mediador del alumnado, convirtiéndose, además, en una potente estrategia de instrucción para la educación inclusiva (Pujolas, 2003).

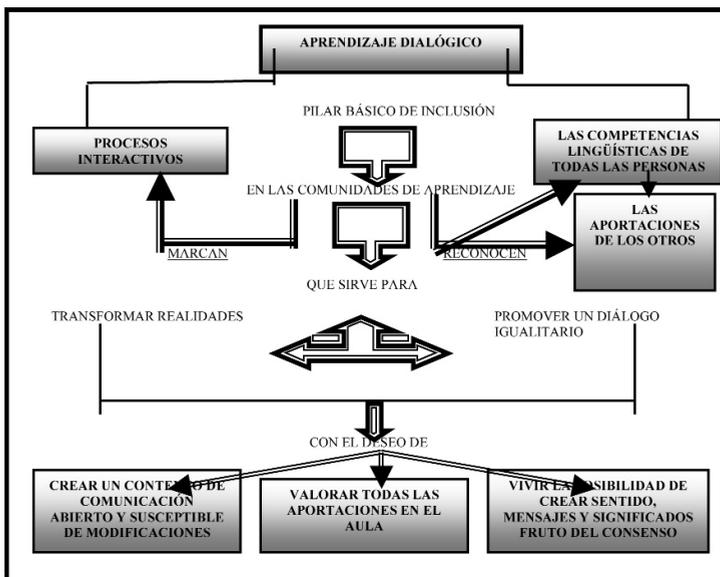
Una de las estrategias que más se han utilizado en las aulas para hacer efectivo el aprendizaje cooperativo ha sido la «tutoría entre iguales». Esta opción permite que los alumnos, en este caso también los alumnos con necesidades educativas especiales, puedan desempeñar el papel de tutores de otros estudiantes y a la vez de estudiantes más jóvenes. Este aprendizaje cooperativo se refiere a situaciones en las que dos o más alumnos interactúan con el objetivo común de dominar el material académico; donde las estrategias de aprendizaje cooperativo son muy diferentes, extendiéndose desde las que se han centrado en situaciones en las que intervienen dos alumnos para representar la forma más simple de cooperación hasta las que han incluido la organización de toda la clase (Pérez Galán, 2003). En definitiva, se trata de apostar por un aprendizaje que tenga en cuenta que los compañeros son un apoyo didáctico permanente, y que el contraste de pareceres entre los niños y la interacción entre compañeros en la misma clase les lleva a enfrentarse a sus propios errores y también a sus propias posibilidades. En este sentido, «las escuelas deben ser organizaciones en las que todos, tanto alumnos como profesores, participen en la tarea de aprender, en un ambiente de cooperación» (Ainscow, M., 1995).

El *aprendizaje «colaborativo» entre docentes* es otro elemento importante en la estructura de una comunidad de aprendizaje. Son muchos los autores que vinculan el aprendizaje en equipo del alumnado a la enseñanza en equipo del profesorado. El uso sistemático del aprendizaje cooperativo necesita de un alto grado de coordinación entre todo el profesorado. Como afirma Sarason (1982), debido al énfasis que se le ha dado a la competitividad, al logro personal, a las destrezas utilitaristas, a la atomización del conocimiento, a la división del trabajo y al relativismo moral oficial, el individualismo se ha convertido en la base de nuestra escuela entretejiéndose persuasivamente en la conciencia de nuestros niños, y también en la de los maestros; es decir, las prácticas escolares tradicionales asumen aquí la calidad de lugares comunes no susceptibles de análisis. Pero el trabajo de colaboración entre docentes, en tanto que modalidad de aprendizaje entre iguales, constituye, además, una oportunidad de formación y perfeccionamiento con independencia de cuál sea el método instructivo utilizado en el aula. Los profesores, al igual que otros profesionales, aprenden a partir de la interacción con otros colegas, y la coordinación entre ellos se convierte, además, en una herramienta imprescindible para favorecer una práctica

reflexiva, así como un recurso esencial para alcanzar la máxima eficacia en su labor docente. Por tanto, de la misma forma que se ha hablado del trabajo cooperativo entre el alumnado como una estrategia esencial de aprendizaje, hay que hablar también de la colaboración docente como un recurso imprescindible para optimizar la enseñanza.

El *aprendizaje dialógico* es otro pilar básico sobre el que se sustenta la comunidad de aprendizaje. Nos dice Valls (2000) que el aprendizaje dialógico es el que resulta de las interacciones que produce el diálogo igualitario, entre iguales, para llegar al consenso con pretensiones de validez. El diálogo nos hace de motor, nos sitúa, va construyendo la narración. No se trata de hablar por hablar. Se trata de conversar, de dialogar, de cooperar intelectualmente, de discutir de manera argumentada, de valorar las diferentes posiciones. Si un término pudiese sintetizar el fundamento de la experiencia de comunidades de aprendizaje, éste sería la palabra *diálogo*, no en el ya anunciado sentido de verborrea o conversación sin una finalidad, sino en un sentido comunicativo (Freire, 1997a). El diálogo es igualitario cuando las aportaciones de cada participante son valoradas en función de sus argumentaciones (pretensiones de validez), y no quien lo dice (pretensiones de poder). Todo esto lo podemos observar en el siguiente cuadro que presentamos, derivado de las pretensiones y deseos del aprendizaje dialógico en un contexto relevante de significados.

GRÁFICO I. Estructura y procesos interactivos del aprendizaje dialógico en la creación y transformación de significados en el aula



En las comunidades de aprendizaje, todas las personas pueden aportar algo al proceso educativo, desde el sueño o la selección de prioridades hasta las propias actividades de aprendizaje. Así lo percibe Ainscow (1995) cuando habla de que hay que aprender a aprender de la diferencia, en cuanto que son personas que pueden aportar ideas y actitudes valiosas al propio acto didáctico y comunicativo. El diálogo tiene como objetivo la planificación, realización, evaluación y mejora de los aprendizajes, constituyendo una red de interacción entre los agentes implicados, de la que el alumnado constituye un nudo fundamental.

La *investigación-acción* es otro elemento imprescindible en la composición de una comunidad de aprendizaje. Habla Guerrero (2003) en un artículo reciente, publicado en el libro *Conocimiento, ética y esperanza*, que versa sobre Educación Especial y los procesos de investigación en el aula, que para Lewin (1947), situándolo como el creador de los procesos de investigación en la acción, esta forma de entender los hechos en situaciones diversas era una actividad que iniciaban algunos grupos o comunidades para modificar algún aspecto que afectara a los miembros de dicha comunidad.

Dice Dewey (1989) que la propia naturaleza de las situaciones que viven los individuos les permite una reacción en cadena de dudas que llevan a un proceso de reflexión tendente a la optimización de la realidad/escenario donde se ha producido dicho proceso de reflexión. La función de este «pensamiento reflexivo» sería pues la de transformar una situación en la que se experimenta oscuridad, duda, conflicto o algún tipo de perturbación, en una situación clara, coherente y armoniosa. La investigación práctica supone, en una comunidad de aprendizaje, analizar la propia práctica educativa, por parte de los docentes, con el fin de repensar nuevas situaciones, alternativas, con un compromiso común y solidario: mejorar y cambiar lo establecido. Siempre desde una perspectiva social y comunitaria. Muchas veces, como sostiene Ainscow (1999), la investigación indica que la mayoría de los centros educativos cuentan con más conocimientos de los que emplean, lo que nos hace pensar que de lo que adolecen la mayoría de los centros es de una falta de impulso de colaboración para hacer un uso adecuado de la pericia y de la creatividad presentes en un determinado contexto.

Un profesor que reflexiona sobre su práctica es un profesor, como nos dice Schön (1992), que supera los límites e inconvenientes de la racionalidad técnica; es decir, es una persona y un profesional que centra su actuación en todos los alumnos, en general, y en los alumnos con necesidades educativas especiales, en particular, implicándose de lleno en la educación de todos ellos, sin temor a la incertidumbre que pueda generar un desconocimiento incipiente de la metodología que hay que aplicar en cada situación. Para la adecuada atención a la diversidad necesitamos profesores cons-

cientes de que ni existe conocimiento profesional para cada caso/problema, ni cada problema tiene una solución concreta (Pérez, 1998).

La heterogeneidad en los espacios escolares ayuda a que seamos más personas y mejores profesionales, ya que estos niños nos exigen nuevas propuestas adaptadas cada vez más a sus posibilidades reales. Así lo manifiesta Rodríguez (2004), cuando argumenta que la diversidad en la escuela se ha convertido en un tema de suma importancia y preocupación entre los educadores. La amplia variedad de alumnos, estilos y ritmos de aprendizaje, diferentes niveles de intereses y motivaciones, han puesto de manifiesto la necesidad de interpretar y definir la escuela tanto desde aspectos organizativos como curriculares. Asumir esta variedad requiere entender qué está ocurriendo en nuestro entorno social y, sobre todo, hasta dónde estamos dispuestos a llegar y a arriesgar con estas personas. En una sociedad plural y en una escuela abierta a la comunidad, una de las exigencias prioritarias es entender que el profesor más que ser un técnico debe ser un intelectual crítico y consciente de las dimensiones éticas de la profesión (Giroux, 1990). Este profesional crítico deberá formar a todos los alumnos no solo en habilidades y contenidos, sino también para ser ciudadanos activos y críticos en una sociedad democrática. Pretendemos que todos nuestros alumnos alcancen la idea de democracia crítica, y también los sujetos de atención a la diversidad. Entendemos que no sólo la escuela debe servirles para adaptarse al sistema sino que también para que produzcan cambios en el propio sistema; solo así puede entenderse la verdadera función educativa de la escuela. Se trata de preparar a todos los estudiantes para vivir y contribuir a una sociedad diversa, así como prepararles para reconocer y trabajar para cambiar las desigualdades sociales y económicas de esa sociedad (Cochran-Smith, 1998).

No nos podemos olvidar de un aspecto prioritario en una comunidad de aprendizaje, el de la organización. No es posible mantener la misma estructura formal y de fondo que mantienen los espacios escolares en la nueva configuración del contexto escolar como «comunidad de aprendizaje».

Organizar, en este contexto de aprendizaje, adquiere un significado que va mucho más allá de las aportaciones de los modelos tecnológicos y de gestión. Esta organización, y la forma de organizar los espacios y el currículo, deben ser sensibles con la diversidad a través de diferentes consideraciones y de un cambio de mentalidad en la estructura y gestión de los centros educativos. Desde el ámbito institucional, la escuela debe persistir en la elaboración y configuración conjunta de un proyecto que se asiente, fundamentalmente, en políticas de inclusión y principios de participación conjunta, donde no tenga cabida la segregación y exclusión. Este proyecto debe surgir

de las inquietudes e intereses de una comunidad profesional y educativa que aliente la necesidad de que todos los alumnos tienen aportaciones valiosas desde un sustrato moral y ético; no solo con ellos mismos sino también con aquéllos con los que comparten conocimientos, inquietudes y rechazos.

Todos los proyectos e intenciones educativas son consecuencia de la máxima participación de las personas que integran la comunidad. Como sostiene Martínez Ortiz (2004), la transformación de la escuela en comunidad de aprendizaje plantea la unión de escuela y entorno como agentes educativos. Para que todo esto pueda llevarse a cabo con unas mínimas garantías de éxito se tienen que poner en marcha dos mecanismos:

- Cambiar la mentalidad de recepción de un servicio público por la mentalidad de protagonismo de la gestión pública.
- Servirse de los planteamientos del aprendizaje dialógico para provocar cambios en las pretensiones de poder por pretensiones de validez en el lenguaje y forma de pensar de los agentes de la comunidad de aprendizaje.

Desde el ámbito curricular, todas las actuaciones tienen que acabar concretándose en el tiempo y en el espacio del aula. El currículo debe tener en cuenta la diversidad de cada alumno y alumna. Uno de los aspectos organizativos que tiene más incidencia en la atención a la diversidad es la forma de agrupamiento del alumnado. Los grupos deben ser heterogéneos, tanto en cuestión de género como en nivel de aprendizaje y origen cultural. Es a través de estos grupos heterogéneos como se refuerza el aprendizaje de todos los niños y todas las niñas sin ningún tipo de discriminación (Puigdemívol, 1998; Elboj, 2005).

Desde el ámbito comunitario, el currículo no se compone solamente de experiencias de aula. Una de las características de las comunidades de aprendizaje y del reto de la inclusión se centra en enriquecer los entornos de aprendizaje, llamado también «proceso de aceleración del aprendizaje». El desarrollo de estos proyectos, que la mayoría de las veces se incluyen en actividades que pertenecen a horarios fuera de clase con personal de fuera y de la comunidad, es la fase de puesta en práctica de las ideas transformadas en propuestas participativas para la acción. Así lo entiende Martínez Muñoz (2004) cuando considera que en esta fase la escuela funciona ya como una comunidad de aprendizaje, en la que, generalmente, el centro escolar está abierto más horas que las ordinarias, en algunas aulas hay más personas adultas que el profesorado, y se desarrollan actividades al final del horario lectivo en las que las familias y el alumnado pueden participar.

## El aprendizaje relevante en un contexto de construcción social del conocimiento

El aprendizaje relevante supone, entre otras cosas, la importancia de la construcción de referentes culturales a partir de la interpretación y reinterpretación subjetiva de los mensajes que se emiten en contextos diferentes de producción de significados. Es el sujeto, en compañía de los otros, los que dotan de sentido a todas aquellas cuestiones importantes que acaecen cerca y fuera de sus vidas. Los individuos y los grupos humanos tienen la compleja pero «ilusionante» tarea de construir de manera permanente el sentido de sus representaciones y la legitimidad de sus prácticas.

Afirma Pérez Gómez (2003) que, para las opciones epistemológicas «constructivistas», el conocimiento es un subproducto no de la mente individual, sino de las relaciones grupales en el contexto de una comunidad que ocupa un espacio y se sitúa en tiempo concreto. El conocimiento es un proceso dialógico de construcción de significados en los espacios vitales de interacción social donde el individuo y la colectividad satisfacen sus necesidades.

En la construcción de este conocimiento relevante es fundamental cómo se agrupa a los estudiantes para la enseñanza y el aprendizaje y para la participación en la toma de decisiones; como también lo son los temas sobre los que se les convoque a participar, y el modo en que la misma participación se promueva.

Como advierte Pérez Gómez (2002), la escuela graduada, organizada de forma inflexible en grupos homogéneos de edad y de capacidades, con un mismo currículo, igual metodología y un mismo ritmo para todos, es una escuela preocupada solo por la transmisión de conocimientos y no por el aprendizaje relevante de todos y cada uno de los diferentes alumnos. Por este motivo, la sociedad del conocimiento requiere una escuela diversificada, flexible y comprensiva, con una metodología sensible a los ritmos de cada individuo, donde lo importante no sea la explicación del profesor, sino el trabajo singular de cada aprendiz (a veces en solitario, pero la mayoría en cooperación), que el docente tiene que aprender a diseñar, estimular, orientar y valorar. Como sostiene Contreras (2003), la escuela debe dejar de ser un lugar para ser enseñado más que para aprender, todo dentro de un orden secuenciado, donde la actividad pasa por variantes de un principio fijo: todos hacen lo mismo, a la vez y de la misma manera, para llegar al mismo sitio. Para escapar de esta situación se requiere que los alumnos en las aulas y en su vida paralela a la escuela reconstruyan su afectividad, su comportamiento; que tengan en los espacios escolares experiencias verdaderamente ricas, que comprueben la posibilidad de vincular sus aprendizajes a su propia realidad. En este

sentido, el aprendizaje escolar no puede quedarse en ser meramente significativo, sino que tiene que ser, además, relevante desde el punto de vista de las opciones educativas que se persiguen con una evidente carga de valor. La escuela tendría que ser concebida como «una vivencia que reconstruye la cultura», donde se empiecen a tener en cuenta aspectos básicos sin los cuales no podría darse esta nueva forma de entender y asimilar la cultura, como son:

- Nueva forma de definir el currículo y qué tipos de currícula son importantes e interesantes para los alumnos.
- Grado de complicidad e implicación del alumnado en la elaboración y configuración de aquellos tópicos que les van a ser útiles en la consolidación de un nuevo marco social, político y cultural.
- Delimitación de la función docente y el papel que debe asumir dentro de una filosofía más participativa, de colaboración y de acción social.
- Ajuste de los marcos de actuación profesional a las nuevas exigencias de una pedagogía y didáctica más críticas, interpretativas y comprometidas socialmente.
- Puesta en práctica de un aprendizaje en el aula que les sirva a los alumnos para entender las claves culturales de la sociedad del conocimiento, y también para que lo puedan aplicar en situaciones diversas y en contextos extraescolares.
- Recolocación del conocimiento relevante en las aulas como prototipo de equidad y justicia social entre los alumnos. Acercar la reconstrucción de la cultura a los que por circunstancias diversas muestran deficitarios puntos de partida en el comienzo de un nuevo proceso de aprendizaje.

## **Los proyectos de trabajo y su concreción en una enseñanza comprensiva que valora el trabajo cooperativo y solidario en el aula**

Sin perder de vista sus orígenes (desde la formulación de su significado con las aportaciones valiosas de Kilpatrick, Dewey y otros), entendemos que los proyectos de trabajo, al otorgarle el adjetivo de «trabajo», adquieren un significado diferente.

Ahora lo que se pretende es poner el énfasis en el proceso, en la tarea de implicar al aprendiz en lo que se investiga y en el esfuerzo cognitivo y emocional que comporta

todo proceso de indagación. Los proyectos de trabajo son como una apelación a la inventiva, a la imaginación y a la aventura de enseñar y aprender. De ahí que Martín (2003) nos hable de que la metodología por proyectos pretenda potenciar experiencias de trabajo en equipo que permitan a los alumnos plantearse e intentar resolver cuestiones que suscitan interés, y hacerlo mediante *procesos de investigación colectiva*.

Se trata de situar al alumnado ante retos cuya solución supone un intenso proceso de aprendizaje en el que se adquieren conocimientos previos y en el que se entrenan capacidades necesarias para el trabajo en equipo -argumentación, perspectiva social, autorregulación, habilidades sociales y toma de decisiones-, entre otras. Asimismo, en el trabajo por proyectos cristalizan valores propios de la experiencia democrática, valores tales como el respeto, la igualdad, la participación, el diálogo o el reconocimiento del otro. Junto a ello, el importante nivel de autonomía de los alumnos en la elección, ejecución y evaluación del proyecto, y el carácter pragmático de la investigación, hacen del trabajo por proyectos una experiencia muy gratificante en el proceso de aprendizaje.

La perspectiva educativa de los proyectos de trabajo no constituye una realización, sino una aspiración. Representa una aspiración que sirve de línea en el horizonte para seguir sosteniendo la propuesta de que la escuela puede ser un espacio favorecedor de la equidad y de la emancipación de los individuos. Algo que solo puede ocurrir si se coloca el aprendizaje en el centro del proceso educativo.

A partir de aquí, estamos en condiciones de apuntar y abordar, conjuntamente con las aportaciones de Henández (1998b), una serie de aspectos que parecen fundamentales para situar los proyectos de trabajo en relación con el debate sobre el papel de la escuela en estos tiempos de mudanza, perplejidad e incertidumbre, y que son los siguientes:

- Un proyecto de trabajo supone una concepción de aprender que tiene en cuenta voces distintas.
- Aprender está relacionado con la elaboración de una conversación cultural.
- El aprendizaje se realiza de manera situada.
- Un proyecto de trabajo podría ser considerado como un formato abierto para la indagación.
- En un proyecto de trabajo no importa de dónde sale el tema o la circunstancia que inicia el proyecto.
- La esencia del proyecto de trabajo descansa sobre la idea de que la cultura escolar debe abrir nuevas puertas al conocimiento.

- Con los proyectos de trabajo se pretende reconocer el camino que va de la información al conocimiento.
- En los proyectos de trabajo se asume una perspectiva multiculturalista.

Las cuestiones más relevantes que se derivan de la utilización práctica y funcional de los proyectos de trabajo no es su instrumentación básica en los procesos de enseñanza-aprendizaje, ni tampoco a la hora de buscar y procesar información, sino la idea de que los proyectos de trabajo no pueden separarse de un sistema social en proceso de cambio.

En el concierto de las sociedades democráticas, se considera que una de las finalidades de los proyectos de trabajo debe ser aprender a tomar decisiones y asumir responsabilidades, y a no dejar que la propia voz sea silenciada por quienes más gustan o proyectan formas de exclusión. Todo esto dentro de un currículo integrado de base transdisciplinar, como resultado de un diálogo permanente con la «realidad» y con los cambios que tienen lugar en la sociedad, el conocimiento y los sujetos pedagógicos.

Desde esta realidad, la tarea del docente frente a la propuesta innovadora de los proyectos de trabajo debe ser la de plantear preguntas que desafíen a los alumnos a examinar sus supuestos actuales, si se considera que son conceptualmente inadecuados. Lo que se pretende, según la aportación que hacen Steinberg y Kincheloe (2000), es la de perseguir una idea de aprendizaje «profundo» que no es solo cognitiva, sino que supone que «cambia la propia identidad», en la medida que compromete nuestro deseo, capta nuestra imaginación y construye nuestra conciencia.

El fin último es implicar al profesorado, las familias, la comunidad y los sujetos infantiles y adolescentes en un múltiple proceso de aprendizaje; un proceso educativo en el que tiene una especial importancia la fascinación, la colaboración, el cuestionamiento, la exploración, el descubrimiento, la creatividad y la reflexión. En definitiva, se habla de lo que Polanyi (1958) denominó con acierto «la apasionada participación en el acto de conocer».

## **Una mirada hacia el futuro: el encuentro entre la escuela inclusiva y la comunidad de aprendizaje**

En este nuevo siglo, observamos un tránsito importante entre la Educación Especial y los procesos de integración a favor de una enseñanza más inclusiva en las comunidades de aprendizaje. En este tiempo, una de las características de los sistemas educativos

mundiales ha sido el análisis y sentido que se le ha dado a la inclusión educativa. Parte de este análisis procede del debate continuo, en el campo de la Educación Especial, sobre cuál es el mejor sistema de provisión de servicios para los alumnos que presentan discapacidad. La inclusión o la educación inclusiva sigue siendo un tema complejo y de frecuente controversia. En los últimos años su significado se ha ampliado notablemente. De acuerdo con Farrell y Ainscow (2002), la inclusión es el proceso a través del cual las escuelas, las comunidades, las autoridades locales y los gobiernos se esfuerzan por reducir las barreras a la participación y al aprendizaje de todos los ciudadanos.

Una escuela inclusiva es aquella que se preocupa de la enseñanza, del aprendizaje, del rendimiento y del bienestar de todos y cada uno de sus alumnos. La educación inclusiva pretende asegurar, además, que todos los alumnos consigan resultados significativos y válidos para sus vidas. Es en la Conferencia Mundial sobre Necesidades Educativas Especiales promovida en 1994, por situar cronológicamente el inicio de una nueva forma de enseñar y contar con todos los alumnos con problemas, donde se establece un marco de acción cuyo principio rector expone que las escuelas deben acoger a todos los niños, independientemente de sus condiciones físicas, intelectuales, sociales, emocionales, lingüísticas u otras. A partir del mismo, las escuelas se enfrentan con el reto de desarrollar una pedagogía capaz de educar con éxito a todos los niños, incluso a aquéllos que sufren discapacidades graves. Además, plantea que las escuelas deben ser «comunidades» que acojan a todos, ya que las diferencias humanas son normales, teniéndose que adaptar el aprendizaje a cada niño (Tierney, 1993).

La educación inclusiva es sinónimo de diversidad, y no un movimiento homogéneo o una simple escuela de pensamiento. Es también un horizonte que, como dirían algunos autores refiriéndose a las «utopías», se mueve de forma constante, precisamente para que no nos paremos o consideremos que ya se ha alcanzado. Con todo, se trata de un proceso o ciclo continuo de reflexión y acción.

La filosofía de la inclusión defiende una educación eficaz para todos, sustentada en que los centros, en tanto comunidades educativas, deben satisfacer las necesidades de todos los alumnos, sean cuales fueren sus características personales, psicológicas o sociales (con independencia de si tienen o no discapacidades). El término «educación inclusiva» es, por tanto, mucho más que un simple acomodo o cambio de terminología para sustituir a la expresión «educación especial», aunque hayan surgido de este ámbito algunos de sus más claros exponentes y numerosas experiencias. Estas últimas han permitido repensar, una y otra vez, la validez de nuestros sistemas escolares en relación con su capacidad para ofrecer una respuesta educativa de calidad para todos. Se trata de algo que tiene que ver, a la larga, con todos los alumnos y con todos los centros

escolares, y que afecta al sistema educativo en su conjunto. Así planteado, se podría hablar simplemente de avanzar hacia una buena educación para todos, siempre que en este término –«todos y todas»– no tuviera cabida ningún tipo de exclusión. Exclusión educativa que se ha caracterizado, en unos casos, por su imposibilidad de acceder a la educación reglada, y en otros por haberlo hecho en condiciones muy precarias o por muy poco tiempo. La educación inclusiva defiende que las escuelas regulares son para todos (Arnáiz, 1996), lo que implica una serie de cambios tanto en la dinámica organizativa como curricular. La escuela debe ser un espacio genuino de convivencia social y comunidad donde, como nos indica Fernando Hernández (2004), ésta se presente más como una cuestión de relación que de saber, y donde la incertidumbre creada excite permanentemente la curiosidad de los alumnos. No hay otra forma de saber más relevante, en cualquier lugar, y más aún en la escuela, que aquélla que provenga de la pedagogía del deseo, ya que establece las condiciones óptimas para aprender sin esfuerzo en compañía de otros. Con la única pretensión de avanzar en el conocimiento a través de la aventura compartida, mediante el cambio sistemático de los contextos escolares, y de la apuesta decidida por la escuela inclusiva en comunidades de aprendizaje. Es en este encuentro en comunidad donde las pequeñas y grandes cosas, y los hechos más o menos relevantes, alcanzan la trascendencia deseada, ya que cada cual aporta lo más valioso y genuino de su pensamiento, de sus acciones y de su forma de ver el mundo.

El dilema que tiene la mayoría de los centros ya no es cómo integrar a algunos alumnos con necesidades educativas especiales, puesto que ya están escolarizados en los mismos, sino cómo desarrollar en ellos un sentido de comunidad y apoyo mutuo que fomente el éxito de todos sus miembros. Uno de los conceptos básicos en las escuelas y clases inclusivas es la construcción de la comunidad. El valor de este concepto consiste en hacer que todos los alumnos se sientan aceptados, valorados, y satisfechas sus necesidades educativas. Pero no solamente esto, sino también conseguir que todos los miembros puedan ayudar y apoyar a los otros, aumentando el «autoconcepto» y el sentido de pertenencia a la comunidad (Arnáiz, 2003). Es en esta idea de considerar al otro donde reside la virtud de la comunidad de aprendizaje.

Se trata de promover modelos organizativos (comunidad de aprendizaje, escuela inclusiva) que favorezcan la aceptación, el apoyo mutuo, a la vez que satisfagan las necesidades educativas. En las comunidades inclusivas los talentos de cada uno, incluyendo los de aquéllos con profundas discapacidades, «superdotación» o comportamientos atípicos, son reconocidos, fomentados y utilizados al máximo. Cada persona es un miembro importante y valioso con responsabilidades y con un papel que desempeñar para apoyar a los demás; lo que fomenta la autoestima, el orgullo en los

logros, el respeto mutuo y un sentido de pertenencia y valía (Lawrence, 1996) . Se trata de una concepción de escuela como ecosistema social, que favorece la idea de comunidad en un contexto social, histórico y organizativo. Se trata de que los centros educativos, como argumenta Gairín (1999), sean considerados como comunidades humanas que funcionan como un todo, donde las relaciones sean de naturaleza circular, dinámica y de proceso; y la diversidad sea responsabilidad de todos, y no de unos pocos. De lo que se trata, fundamentalmente, en este encuentro entre Escuela Inclusiva y Comunidad de Aprendizaje es de lo siguiente:

- de establecer una educación de calidad para todos los alumnos;
- de abandonar la idea de emplazamiento físico por la de participación plena en la vida de la escuela y de la comunidad;
- de que la escuela, el centro y el aula se organicen de manera que satisfagan todas las necesidades;
- y de que los apoyos se dirijan a todos los alumnos.

Para que se cumpla todo esto hace falta algo más que las rectificaciones sobre el currículo y las formas de organizar los espacios escolares. Tampoco nos vale la idea de legislar mucho para cambiar poco. Habrá que creer en la posibilidad del trabajo profesional de los maestros en colaboración con otros, en su capacidad de decisión y, sobre todo, en el interés por cambiar y mejorar, permanentemente, su práctica profesional. En definitiva, creer en otra escuela.

## Conclusión

Consideramos que la escuela debe ser un espacio de comprensión, inclusión y colaboración conjunta. Todos los individuos deben apropiarse de una cultura común que se adapte a las posibilidades y necesidades de cada uno. Hemos hecho hincapié en que la mejor forma de apoderarse de esos referentes culturales es a través de la ayuda mutua y de la actuación conjunta.

Hemos visto cómo este tipo de escuela inclusiva, comprensiva, es respetuosa con todo el mundo y presta la ayuda necesaria a cada individuo, sin ningún tipo de exclusión y de injusticia social. La escuela inclusiva y las comunidades de aprendizaje favo-

recen procesos de recreación cultural y de intercambio comunicativo porque tienen en cuenta las aportaciones de todos. Tienen en cuenta las diferencias y promueven procesos de construcción de significados en el aula. Pensamos que los maestros, dentro de esta nueva realidad, deben adoptar una actitud diferente, deben cambiar sus prácticas educativas y generar un conocimiento más plural y diverso.

La inclusión, las comunidades de aprendizaje y los proyectos de trabajo favorecen el intercambio de significados dentro y fuera del aula porque respetan a cada individuo, en función de cómo es y de la necesidad que tiene. Solo un aprendizaje basado en el diálogo y en la comprensión del otro provoca un auténtico aprendizaje relevante. Solo una forma distinta de aproximarnos a los espacios escolares, y a las vivencias que proyectan en ellos las conciencias de los individuos, nos hacen entender y comprender mejor su significado.

Creemos que el conocimiento debe ser una herramienta útil y valiosa en manos de todo el mundo, sin exclusiones; y que la mejor forma de adquirirlo es a través del otro y con el otro, en una labor conjunta. Confiamos en una escuela diversa que tenga en cuenta las aportaciones de cada uno, y las considere siempre. La finalidad de esta nueva forma de entender lo que sucede en las aulas y cómo debe suceder, después del análisis realizado, está en la apuesta que cada uno hace con las vidas que maneja, y cómo interpreta formas diversas de hacer y trabajar en la escuela (proyectos de trabajo, participación democrática, colaboración permanente, apoyo continuo al otro, discusión compartida, etc.); aunque sabemos que la realidad y su conocimiento debe ser distinto, según cada sujeto en particular. Sabemos, y así lo hemos expuesto en este artículo, que la formulación y recreación de situaciones previamente consideradas, contrastadas y fruto de una reflexión compartida (investigación-acción, aprendizaje «colaborativo», aprendizaje democrático, aula como espacio de vivencia cultural, entornos enriquecidos, etc.), ocasionan efectos positivos en los individuos que aprenden, en nuestras vidas y en nuestra forma de enseñar.

## Referencias bibliográficas

- AINSCOW, M. (1995). *Necesidades especiales en el aula. Guía para la formación del profesorado*. Madrid: Narcea.
- (1999). *Desarrollo de Escuelas inclusivas. Ideas, propuestas y experiencias para mejorar las instituciones escolares*. Madrid: Narcea.

- ARNÁIZ SÁNCHEZ, P. (1996). Las escuelas son para todos. *Siglo Cero*, 27 (2), 25-34.
- (2003) *Educación inclusiva: una escuela para todos*. Málaga: Aljibe.
- CONTRERAS DOMINGO, J. (2003), *Más allá del Academicismo. Los desafíos de la escuela en la era de la información y de la perplejidad*. Málaga: Universidad de Málaga.
- COHRAN-SMITH, M. (1998). Teacher Development and Education Reform. En A. Heargreaves et al. (eds.) *International Handbook of Education Change*. Londres: Kluwer.
- DARLING-HAMMOND, L. (2001). Hay otras escuelas. *Revista Kkikirikí*, 70, 39-43.
- DEWEY, J. C. (1989). *Cómo Pensamos*. Barcelona: Paidós.
- ELBOJ, C., PUIGDELLÍVOL, I., SOLER, M. Y VALLS, R. (2002). *Comunidades de aprendizaje. Transformar la educación*. Barcelona: Graó.
- ELBOJ SASO, C. (2005). *Comunidades de aprendizaje: educar desde la igualdad de diferencias*. Zaragoza: Gobierno de Aragón.
- FARRELL, P. & AINSCOW, M. (2002). *Make special education inclusive*. London: David Fulton Publishers.
- FLECHA, R. (1997). *Compartiendo palabras. El aprendizaje de las personas adultas a través del diálogo*. Barcelona: Paidós.
- FREIRE, P. (1997a). *A la sombra de este árbol*. Barcelona: El Roure.
- (1997b). *Pedagogía de la autonomía*. Madrid: Siglo XXI.
- GAIRÍN, J. (1999). Los departamentos y equipos educativos como órganos de desarrollo profesional. *Revista de Currículo y Formación del Profesorado*, 3 (1), 85-114.
- GARDNER, H. (1988). *La educación de la mente y el conocimiento de las disciplinas*. Barcelona: Paidós.
- GIROUX, H. (1990). *Los profesores como intelectuales: hacia una pedagogía crítica del aprendizaje*. Barcelona: Paidós.
- GUERRERO LÓPEZ, J. F. (2003). La búsqueda de la generalización-estandarización en los modelos de intervención en educación especial: recetas mágicas versus investigación acción. En M. SANTOS GUERRA Y R. BELTRÁN DUARTE. *Conocimiento, Ética y Esperanza*. Málaga: Universidad de Málaga.
- HERNÁNDEZ, F. (1998b). Os proyectos de trabalho e a necessidade de transformar a escola (I). *Prensa Pedagógica*, 20, 53-60.
- (2004). La incertidumbre excita la curiosidad. *Cuadernos de Pedagogía*, Noviembre, 340, 46-51.
- JOHNSON, D., JOHNSON, R. Y HOLUBEC, E. (1999). *El aprendizaje cooperativo en el aula*. Barcelona: Paidós.

- LAWRENCE, D. (1996). *Enhancing Self-Esteem in the Classroom*. London: Paul Chapman Publishing.
- LEWIN, K. (1947). *Field theory in social science*. New York. Harger and Row. Citado por ELLIOT (1990), op.cit.
- MARÍN, N. Y SOLER, M. (2004). Una comunidad de aprendizaje. *Cuadernos de Pedagogía*, 331, 60-62.
- MARTÍN GARCÍA, X. (2003). Proyectos de trabajo. *Cuadernos de Pedagogía*, Junio, 325, 76-77
- MARTÍNEZ MUÑOZ, M. (2004). El desarrollo de una comunidad de aprendizaje. *Temáticos Escuela*, 11, 11-14.
- MARTÍNEZ ORTIZ, S. (2004). Aspectos organizativos de una comunidad de aprendizaje. *Temáticos Escuela*, 11, 15-18.
- MERIEU, P. H. (1998). *Frankenstein educador*. Barcelona: Alertes.
- MONEREO, C. Y DURÁN, D. (2002). *Entramados. Métodos de aprendizaje cooperativo y «colaborativo»*. Barcelona: Edebé.
- PÉREZ GALÁN, R. (2003). *Educación Especial, Atención a la diversidad y LOCE: un enfoque teórico y comprensivo*. Málaga: Aljibe.
- PÉREZ GÓMEZ, A. I. (1998). *La cultura escolar en la sociedad neoliberal*. Madrid: Morata.
- (1998). El pensamiento práctico del profesor: implicaciones para la formación del profesorado. En A. VILLA, *Perspectivas y problemas de la función docente*. Madrid: Narcea.
- (2002). Un aprendizaje diverso y relevante. *Cuadernos de Pedagogía*, Marzo, 311, 66-70.
- Aprender a ser, aprender a vivir. A favor de la diversidad en la escuela educativa. En M. A. SANTOS GUERRA, M. A. y R. BELTRÁN DUARTE, R. (2003): *Conocimiento, Ética y Esperanza*. Málaga: Universidad de Málaga.
- POLANYI, M. (1958). *Personal Knowledge: towards a post-critical philosophy*. London: Routledge and Kegan Paul.
- PUIGDELLÍVOL, I. (1998) *La educación especial en la escuela integrada: una perspectiva desde la diversidad*. Barcelona: Graó.
- PUJOLÁS, P. (2003). *Aprender junts alumnes diferents. Els equips d'aprenentatge cooperatiu a l'aula*. Vic: Eumo Editorial.
- RODRÍGUEZ IZQUIERDO, R. (2004). Atención a la diversidad cultural en la escuela. Propuestas de intervención socioeducativa. *Educación y futuro*, 10, 37-47.
- SARASON, S. (1982). *The culture of the school and the problem of change*. Boston: Allyn y Bacon.

- SCHÖN, D.A. (1992). *La formación de profesionales reflexivos*. Madrid: Piados-MEC.
- SLAVING, R. (1985). *Learning to cooperate, cooperating to learn*. New York: Plenum Press.
- STEINBERG, S. H. Y KINCHELOEVE, L. (2000). *Cultura infantil y multinacionales*. Madrid: Morata.
- TEDESCO, J. C. (2000). Educación y sociedad del conocimiento. *Cuadernos de pedagogía*, 288, 82-86.
- TIERNEY, W. G. (1993). *Building Communities of Difference. Higher Education in the Twenty-First Century*. London: Bergin & Garvey.
- VALLS CAROL, R. (2000). *Las comunidades de aprendizaje. Una práctica educativa de aprendizaje dialógico para la sociedad de la información*. Tesis doctoral inédita. Barcelona: Universidad de Barcelona
- WELLS, G. & CLAXTON, G. (2002). *Learning for life in the 21 st Century*. New York: Teachers College.

**Dirección de contacto:** Rafael Pérez Galán. Universidad de Málaga. Escuela Universitaria de Magisterio María Inmaculada. c/ Presbiterio Vidaurreta, s/n, 29200. Antequera, Málaga, España.  
e-mail: rafaperezga@hotmail.com

# Mitos y realidades en la carrera docente

## Myths and realities of a career in teaching

Encarnación Sánchez Lissen

*Universidad de Sevilla. Departamento de Teoría e Historia de la Educación y Pedagogía Social. Sevilla, España*

«En el próximo decenio será necesario contratar  
a más de 18 millones de docentes en todo el mundo»

*UNESCO, 2006*

### Resumen

A lo largo de este artículo, hemos querido presentar algunos de los factores que más han influido y han condicionando a los jóvenes en la elección de la carrera de magisterio. Comenzamos valorando esta profesión y señalando el reconocimiento que hace la sociedad de ella, así como algunas consideraciones que han sido emitidas por los organismos internacionales.

Dado que la enseñanza ha estado ancestralmente vinculada al valor de la vocación, es bastante habitual reconocer que es éste y no otro, el motivo principal que estimula a los aspirantes a maestro; sin embargo, se trata de un testimonio que conviene matizar ya que un amplio número de investigaciones (De la Rosa Acosta, B., 1969; Ortega, F y Varela, J., 1985; González Sanmamed, M., 1995; Sánchez Lissen, E., 2002) entre otras, coinciden en reconocer que coexisten junto a la vocación otros factores como: el salario, tratarse de una carrera corta, o el número de vacaciones de las que disfruta un docente y que son tan determinantes como aquella en el proceso de elección de una carrera. Todos ellos son motivaciones -posiblemente ni suficientes ni necesarias- pero motivos al fin y al cabo, que les llevan a inclinarse por esta carrera y por esta profesión.

Analizamos algunos de estos elementos y comprobamos cuánto hay de mito y cuánto de realidad en cada uno de ellos.

*Palabras clave:* magisterio, profesión, valoración social, vocación, motivaciones, elección y selección.

### **Abstract**

This paper aims to set out some of the most influential and determining factors for young people to choose a university degree in teaching. We start by assessing this career and pointing out the appreciation that society shows for it, as well as some considerations issued by International Organizations.

As teaching has long been linked to the notion of vocation, we tend to think that this, and little else, is the main motive that encourages people to become schoolteachers. However, this statement has to be qualified, since there are many research studies (Rosa Acosta, B., 1969; Ortega, F and Varela, J., 1985; González Sanmamed, M., 1995; Sánchez Lissen, E., 2002), among others, which indicate that, along with vocation, other factors such as wage, the short three-year degree course or the long leave schoolteachers have, also attract young people to choose a university degree in teaching. These motives - which may not be enough or necessary, but are motives just the same - may lead students to take up this degree and subsequent career.

This paper analyses some of these suggested motivations and attempts to separate the element of myth and of reality in each of them.

*Key Words:* teaching degree, career, social assessment, vocation, motivations, choice and selection.

## **Introducción**

Tras conocer los datos publicados en el año 2006 por el Instituto de Estadística de la UNESCO (IEU), cabría preguntarse ¿pero sobran o faltan profesores? No sé si podremos dar una respuesta concluyente a esta pregunta, pero al menos, intentaremos descubrir si están todos los que son en esta profesión, si aún es una carrera recurrente para un amplio número de estudiantes y en su caso, qué aspectos les atraen de ella.

Bien es cierto que encontramos casi a diario en los medios de comunicación, noticias que reflejan situaciones de conflicto y de violencia en los centros educativos de nuestro país; junto a ellos, se siguen publicando informes nacionales que detallan los problemas de estrés, salud mental, etc., que padecen muchos docentes en estos momentos. Este cúmulo de circunstancias nos podría hacer sospechar que la enseñanza se encuentra en declive, que no pasa por sus mejores momentos y que por ella se interesan cada vez menos jóvenes; sin embargo, no deberíamos ser tan categóricos porque no parece ser ésta la única versión que existe en torno al tema y de hecho, varios argumentos que se exponen en este artículo reconocerán otro lado menos amargo de la profesión.

Ciertamente, la carrera docente es una profesión valiosa pero no siempre valorada justamente; una circunstancia en la que influye su consideración social y que incide en los jóvenes para su elección. Quizá todos participamos en este vago reconocimiento y todos tenemos un poco de culpa; unos, los que dirigen y tienen capacidad de decisión política por no reconocer el valor de esta práctica y otros, los propios docentes, por no defender con más empeño o favorecer nuestra propia tarea<sup>1</sup>.

A pesar de estos contrastes y del amplio abanico de circunstancias que envuelven a esta profesión, existen distintos argumentos que nos hacen pensar que en pleno siglo XXI, la carrera docente es aún atractiva para muchos jóvenes. Una muestra de ello es que las Escuelas de Magisterio siguen cubriendo todas las plazas cada nuevo curso y aunque desde hace algunos años, la enseñanza universitaria ha ido sufriendo un descenso considerable en el número de estudiantes, sin embargo, éste no se acentúa en el Magisterio más allá que en otras carreras<sup>2</sup>.

No obstante, el que la carrera docente se encuentre sometida a fuertes presiones condiciona y frena a muchos jóvenes su deseo de acceder al magisterio. Esta circunstancia está más acentuada en unos países que en otros; en Suecia, por ejemplo, declaran tener en los últimos años ciertos problemas para reclutar nuevos docentes (Esteve, Franco y Vera, 1995), un indicador que se puede extender al resto de países del mundo y que, en breve, puede afectar al proceso de contratación de nuevos docentes. Esta misma apreciación se pone de manifiesto en uno de los últimos estudios de la OECD (Riobó, 2003).

Otro caso es el Reino Unido, donde el titular del periódico *El País*<sup>3</sup> ponía de manifiesto este temor: «La falta de profesores lastra el inicio del curso escolar en el Reino Unido».

Una muestra más de la preocupación que existe para la contratación de nuevo profesorado, la podemos encontrar en el lema que marcó la conmemoración del Día Mundial del Docente el año 2004; éste fue el siguiente: «Objetivo: Contratación y retención de docentes de calidad».

Como decía anteriormente, parece que nuestro país aún se encuentra con cierto desahogo en este sentido, dado que muchos jóvenes siguen apostando por la carrera

<sup>(1)</sup> A nivel internacional, el estudio de Eurydice (2004b) recoge una amplia muestra sobre las consideraciones que la opinión pública hace de los docentes, lo que dice de ellos, así como lo que la sociedad declara al respecto. Básicamente, el estudio señala que «la profesión docente goza de mejor imagen de lo que el profesorado piensa». Se pueden consultar los detalles de este trabajo en la siguiente dirección: <http://www.eurydice.org/Documents/KeyTopics3/es/KT3R4ES.pdf>

<sup>(2)</sup> Como ejemplo, valga el caso de la Universidad de Sevilla (Anuario Estadístico 2004-05, p.176) donde los datos acumulados en el trienio 2002-05 reflejan que tan solo las especialidades de Educación Física y Lengua Extranjera han experimentado un leve descenso en el número de matriculados en el último curso académico (-2,78% y 2,61%, respectivamente); el resto ha sufrido un ligero aumento.

<sup>(3)</sup> La noticia apareció en *El País*, con fecha de 11 de septiembre de 2000

docente; sin embargo, también parece evidente que a la vista de los datos publicados sobre las dificultades que existen en otros países para la contratación de nuevos profesores<sup>4</sup>, este tema comienza a ser preocupante.

Ante esta breve presentación en la que se mezclan deseos y decepciones, trataremos de exponer cuáles son los motivos que lleva a los estudiantes a inclinarse, en estos momentos, por esta carrera. Entre las muchas motivaciones que tienen los aspirantes a maestro, analizaremos algunos de los mitos y realidades que más influyen en la elección de la carrera. Con todo ello, podremos conocer algo más del interior y el exterior del profesor, de sus estímulos, de las aspiraciones y del conjunto de factores que rodean este proceso.

## La enseñanza: una labor controvertible

Muchos convendrán conmigo que tanto en la enseñanza como en el conjunto de quehaceres vinculados al sistema educativo de un país, existen dos figuras esenciales como son el alumno y el profesor. Está claro que tanto el uno como el otro son importantes, son necesarios, aunque a veces nos hayamos preguntado si la enseñanza es posible sin el docente. Hoy por hoy, algunos nos atrevemos a decir rotundamente que no. En estos momentos, aunque las Nuevas Tecnologías parezcan ser su gran competidor, sin embargo, nada hace peligrar su presencia; es más, cabría recordar que se siguen necesitando profesores con ilusiones y ganas de trabajar. Puede parecer insólito que, a pesar de los inconvenientes con los que se viene encontrando esta profesión, sin embargo, haya un número importante de estudiantes que, en primera opción, siguen acudiendo entusiasmados a las Escuelas de Magisterio. Como sabemos, la docencia es una actividad socialmente acreditada y por tanto, no es difícil encontrar argumentos que justifiquen la importancia de su labor ni de la propia figura personal. Los documentos emitidos por organismos o conferencias internacionales avalan esta versión. Un ejemplo es el Informe de la OIT UNESCO, del 5 de octubre de 1966, que incluye la Recomendación relativa al docente y que fue aprobada en París por la Conferencia Intergubernamental Especial sobre la Situación del Personal Docente. En este

---

<sup>4)</sup> Es significativo el Informe elaborado en el año 2006 por el Instituto de Estadística de la UNESCO (IEU), en el que se incluyen datos preocupantes sobre la situación precaria de los docentes de muchos países y la necesidad que tienen otros, de reclutar más docentes. Uno de ellos es el África Subsahariana que necesitará 1.600.000 maestros más en el 2015 (UNESCO: Comunicado de prensa No 2006-34).

documento se señala expresamente la importancia de este colectivo y «se reconoce el papel esencial del personal docente en el progreso de la educación y la importancia de su contribución al desarrollo de la personalidad humana y de la sociedad moderna». Ya en aquel año, el papel atribuido a los profesionales de la enseñanza comenzaba a quedar patente en los informes internacionales. Ciertamente no se trata de un asunto localista, sino que el tema tiene un tratamiento universal, tal como lo es la propia figura a quién nos referimos.

Más recientemente, el Informe de 2005 *Educación para Todos: el imperativo de la calidad*, que elabora cada año la UNESCO se centra en esta ocasión en el valor de la calidad y reconoce en este logro, la importancia del docente; esto supone que se llegue a catalogar como uno de los colectivos «que más influencia tiene en el aprendizaje de los niños».

Estos documentos como otros tantos, nos ofrecen datos muy evidentes del valor de la profesión y de la valía del que la ejerce. En esencia, la calidad va inexorablemente unida al docente y a sus funciones; es el fin y son los medios y por tanto, en ello radica buena parte de la carrera docente.

## Los aspirantes a maestro: entre la selección, orientación y motivación

Pero, ¿qué motivación puede tener un estudiante en estos momentos para elegir la carrera de magisterio?

Si echamos una mirada a los medios de comunicación, podemos encontrar noticias verdaderamente alarmantes que quitan las ganas a cualquiera para acceder a esta profesión. Valgan como ejemplo las dos siguientes:

- «Los docentes de Granada secundan el tercer paro en 15 días por la violencia en las aulas andaluzas» (El País, 23-3-2006).
- «Dos menores propinan una paliza a un profesor granadino durante las fiestas de un instituto del poniente almeriense» (Granada Digital, 6-3-2006).

Por otra parte, los trabajos de Esteve (1994) sobre el *malestar docente*, demuestran también las dificultades que rodean a este colectivo y las incidencias que sobre la

enseñanza tienen algunos elementos sociales, económicos o políticos entre otros, para el trabajo diario de un profesor.

Pues bien, aunque parece indiscutible que la profesión docente se encuentra actualmente rodeada de múltiples dificultades, sin embargo, sigue siendo considerable el número de aspirantes a maestro que vienen sucediéndose cada nuevo curso académico.

¿Qué motivaciones tienen para ello? ¿Cuánto afecta al ámbito interno y cuánto al ámbito externo de la persona?

En general, entendemos que se suceden en el docente tres momentos reveladores de su motivación; y son: *antes* de acceder a la carrera, *durante* la misma y *después* de finalizar los estudios y comenzar la experiencia profesional. En cada una de estas etapas, la formación experimenta un papel esencial; en el primer caso –el antes–, la formación se presenta como orientación, en el segundo –el durante–, la formación tiene un carácter preparatorio dado que ayuda a despertar los conocimientos, habilidades, destrezas o vocación necesaria para fortalecerse. En este sentido, hemos podido comprobar por algunas de nuestras investigaciones (Sánchez Lissen, 2002, 2003) que las Escuelas suelen ser un detonante a favor del interés y del entusiasmo por esta carrera y, especialmente, por el desarrollo de esta profesión. Por último, el después, es el tercer momento y en él la formación adquiere un carácter de revalidación que ayuda a generar nuevos aprendizajes y nuevas ilusiones.

Si reconocemos que la formación impartida en las Escuelas de Magisterio es importante, no lo es menos la forma de acceso a las mismas y las motivaciones que se generan en torno a ese momento. De hecho, para comprender el conjunto de la enseñanza de un país merece la pena detenernos y analizar el periodo previo al reclutamiento, momento en el que los estudiantes merodean –a veces solo con el pensamiento– hasta llegar a la carrera preferida. Con todo ello, cabría considerar que los elementos que más influyen en el aspirante a maestro están relacionados con sus motivaciones personales o con factores económicos, sociales, académicos y vocacionales.

En este mismo sentido, Mariano Fernández Enguita (2005) alude a la «importancia del reclutamiento, y a mejorar las condiciones de trabajo para atraer a mejores profesores».

Si convenimos en reconocer que el reclutamiento es importante, debemos saber que esta circunstancia no es igual en todos los países y que, de entrada, este aspecto resulta significativo para la consideración que se tiene en cada país de la carrera docente. Cabe señalar el caso de Finlandia. Un ejemplo de país paradigmático en lo que respecta a la consideración que tiene la formación inicial del profesorado, así como la consideración social de los profesionales de la enseñanza.

Ciertamente, al profesorado hay que cuidarlo de manera muy especial, hay que *seleccionarlo* correctamente y además, hay que procurarle la formación tanto inicial como permanente más oportuna. En Inglaterra, organismos como el Teacher Training Agency (TTA) tiene entre sus cometidos procurar la calidad de la formación docente, así como el reclutar estudiantes brillantes para la enseñanza. En la búsqueda de una mayor calidad a la que aspira la educación en general y el sistema escolar en particular, nos parece que el profesorado es el colectivo más determinante en todo ello y en general, en el buen funcionamiento de todo el sistema. Estoy convencida de que, mejorando y mimando este proceso formativo, también podrán mejorar los resultados de su profesión.

## Las principales motivaciones de los estudiantes de magisterio

A finales de los años sesenta, Bernardo de la Rosa realizó un estudio de carácter nacional para conocer cuáles eran los principales factores que incidían entre los estudiantes de magisterio para decantarse por esa carrera. Posteriormente, algunos seguimos sus pasos y con muestras de corte regional o provincial (García Ortiz, 1983; Sánchez Lissen, 1999; Lagares, 2000)<sup>5</sup>, pudimos comprobar que muchos de aquellos factores continuaban muy presentes entre los candidatos algunos años después. Ciertamente, los cambios que se han producido en la sociedad entre una y otras investigaciones no han sido baladí; se han modificado las circunstancias políticas, pero también las económicas y académicas y éstas, si caben, son las que más se han dejado notar en los últimos resultados obtenidos.

A lo largo de estas investigaciones se ha comprobado que muchos aspirantes a Magisterio la han elegido por tratarse de una carrera corta. Esta circunstancia la han aprovechado en muchos casos como trampolín para otra carrera de segundo ciclo o bien, para trabajar de manera inmediata. Lo veremos más adelante y comprobaremos además, que no se trata de una versión exclusiva de los estudios anteriormente citados, sino que lo son igualmente en los trabajos de Félix Ortega y Julia Varela, (1985) o de González Sanmamed (1995) entre otros.

<sup>5</sup> Estas tres Tesis Doctorales han analizado desde distintos órdenes, los factores que han incidido en la elección de la carrera de Magisterio. La tesis de García Ortiz analizó en 1983 a los estudiantes de la Universidad de Sevilla; la profesora Sánchez Lissen (1999), estudió comparativamente en Sevilla, los estudiantes matriculados en esta Universidad y en la Escuela Cardenal Spínola -CEU-, centro adscrito a la misma. Finalmente, el profesor Lagares (2000) se detuvo en los estudiantes onubenses de Magisterio.

De todos ellos y del análisis extraído de los informes nacionales e internacionales elaborados sobre el docente en las tres últimas décadas, mencionamos a continuación cinco factores que entendemos son los más reiterativos y reveladores de los resultados que, habitualmente, se han obtenido. Estos son: el salario, las vacaciones, la duración de la carrera, los exiguos requisitos para el acceso y la vocación. La mayoría de ellos los podemos catalogar como mitos en la elección de la carrera docente y frente a los mitos, otras realidades han marcado el desarrollo de la misma. Para valorar cada uno de estos factores que inciden entre los estudiantes y que le motivan para su acceso, expondremos también las circunstancias que prevalecen en otros países europeos en relación a la carrera docente. Como veremos, muchos de los datos son tremendamente significativos y nos pueden ayudar a entender por qué es distinta la valoración que tienen unos y otros países sobre la profesión docente.

Para este análisis comparativo hemos analizado los diversos países europeos que han formado parte en el año 2003 del Programa de la OCDE para la Evaluación Internacional de los Alumnos (PISA); una circunstancia que nos anticipa datos relevantes sobre el nivel de sus estudiantes y de los que se pueden extrapolar ciertas motivaciones en relación al profesorado. A su vez, conviene en ellos la pertenencia a la Unión Europea de los 27 miembros. De todos hemos seleccionado para este artículo, los que presentan una mayor significatividad en cada uno de los factores enunciados. Con estas dos premisas, las tablas comparativas que proponemos difieren unas de otras en relación a los países propuestos y se incluyen en cada caso una muestra de seis u ocho países que son, en definitiva, los que merecen ser reseñados.

## El salario

Aunque no todos pueden estar de acuerdo conmigo, creo que en estos momentos, el salario de los docentes de primaria es aceptable. Para poder realizar esta afirmación podemos analizar los datos desde distintos frentes. Así, si lo comparamos por ejemplo:

- *Con épocas pasadas.* Es evidente que ya ha quedado añeja aquella famosa frase de «pasa más hambre que un maestro». Aserto que hoy en nada se parece a la realidad y que tanto ha dañado a esta profesión. En una publicación reciente, con motivo del homenaje al maestro en la región de Murcia, Aznar Martínez

(2005)<sup>6</sup> cuenta su experiencia en los años sesenta como maestro en diversas escuelas y recuerda que su sueldo anual rondaba las 17.000 pesetas. Ciertamente, los sueldos han ido mejorando considerablemente y se han adecuados a la realidad de cada momento. Ahora nos parece inconcebible la penuria que ha rodeado a estos profesionales y estamos convencidos que si se mejora su situación económica, se dignifica a la persona y a la profesión.

- *Con otros trabajos.* En este sentido, por el número de horas presenciales de trabajo en la escuela, el sueldo de los maestros puede estar en sintonía con aquellos otros profesionales con igual titulación. Sin embargo, como todos sabemos, el docente desempeña también una labor «en cubierta» donde el número de horas no se contabiliza en ningún caso. Desde este punto de vista se podría interpretar que el sueldo del profesor no es coherente con el número de horas que dedica a la profesión. En general, podemos encontrar en este colectivo un amplio número de parámetros que nos impiden valorar equitativamente a unas y otras ocupaciones.
- *En relación con otros países de la Unión Europea.* Entre los países de la Unión Europea, los datos publicados en el estudio de la OECD (2005) vienen a calificar de muy favorables los sueldos de los docentes de Alemania y Luxemburgo. Por el contrario, países como Irlanda o Suecia ocupan lugares muy deficientes en materia salarial. En general, nuestros docentes se encuentran también entre los mejores pagados. En España, Alemania y Dinamarca, el salario del profesorado novato es superior a la media europea en un 20%. Por tanto, el mayor sueldo para los alemanes y el menor para los docentes de Grecia y Portugal. Sin embargo, se ha comprobado que en este último país se logra uno de los mayores salarios en sus últimos años de profesión (Campuzano, 2005). Aun contando con un sueldo muy aceptable, las diferencias que mantienen países como España y Finlandia son bastante significativas; si en nuestro país un docente puede tener una media mensual de 1800, en Finlandia el importe total puede ascender a 3400 €.

Junto al salario base, existen algunos países en los que se gratifica a sus profesionales con incentivos económicos por ejercer una responsabilidad adicional. Uno de estos casos es Dinamarca o España. Asimismo, en países como EEUU, en algunas escuelas situadas en lugares conflictivos y arriesgados de Chicago, los profesores también tienen un sobresueldo.

<sup>6</sup> El profesor Aznar Martínez rememora que su sueldo anual era de 16.920 pesetas, incluidas las pagas extraordinarias. Aznar Martínez, A. (2005). Homenaje al maestro. En Jornada de Homenaje al Maestro del Consejo Escolar de la Región de Murcia. Murcia, Consejería de Educación y Cultura, Consejo Escolar de la Región de Murcia. 103-108.

Pues bien, ante esta concreción de cifras podríamos afirmar que para los aspirantes a maestro, el sueldo es: ¿un mito o una realidad?

## Las vacaciones

Con respecto a las vacaciones de los profesores, ¿qué hay de verdad y de mentira en este indicador? ¿Realmente, hay tantas fiestas como se le atribuyen?

Aunque se trata de un factor controvertido, sobre todo de cara a la imagen que tiene en la sociedad nuestra profesión, sin embargo, hay que valorar en su justa medida este elemento y reconocer que en muchos casos se ha presentado como un elemento trasgresor de la propia imagen de la profesión. Posiblemente ha jugado en su contra y le ha perjudicado más que beneficiado e incluso se ha vinculado con una *imagen idealizada de la profesión*. En cualquier caso hay que reconocer que tanto las vacaciones –su número y su ubicación en el calendario–, como el número de días lectivos difieren entre los países. Dado que es habitual escuchar que los profesores tienen muchas vacaciones, vamos a ver en el siguiente cuadro las fiestas que se ofrecen en distintos países de la Unión Europea durante del curso escolar. Aunque a su vez, existen algunas oscilaciones en el número de días señalados por cada país, sin embargo y de manera general, podemos comprobar la distribución y especificación de los mismos.

TABLA I. Distribución de vacaciones en varios países de la UE

Distribución de las vacaciones en varios países de la UE						
	Verano	Otoño	Navidad	Primavera/Semana Santa	3º Trimestre	Días festivos
Alemania	6 sem.	2-11 días	1-3 sem.	5-15 días	0-9 días	0-10 días
Bélgica	9 sem.	1 sem.	2 sem.	2 sem.	2 días	5 días
Dinamarca	7 sem.	1 sem.	2 sem.	1 sem.	---	4 días
España	11-12 sem.	---	3 sem.	10 días	---	Unos 7 días
Finlandia	10-12 sem.	1-7 días	10 días (mínimo)	4 días	---	3 días
Francia	9 sem.	1 sem.	2 sem.	2 sem.	---	Algunos días entre Semana Santa y verano
Irlanda	9-13 sem.	1 sem.	2 sem.	2 sem.	---	---
Italia	9 sem.	---	2 sem.	1 sem.	---	5 días
Portugal	11 sem.	1 sem.	2 sem.	2 sem.	---	5 días

Fuente: Eurydice, 2001

Aunque no lo hemos incluido en la tabla anterior, debemos tener en cuenta que algunos países también tienen vacaciones durante las fiestas de Carnaval. Ésta es muy desigual en el número de días que se disfruta en cada país; por ejemplo, en algunas

comunidades autónomas de nuestro país pueden llegar a contar con dos o tres días de fiesta, pero aumenta de manera considerable en el caso de Alemania, Bélgica o Finlandia que acuerdan entre una semana y algo más de 10 días.

Si sumamos el número de días de fiesta concertado en cada país, comprobamos que en España quedan estipuladas entre 17 y 18 semanas de vacaciones, número parecido al que disfrutan en Portugal e Irlanda y algo superior al que se ofrece en Alemania y Bélgica (entre 14 y 15 semanas), en Finlandia (15 semanas), o en Italia (entre 13 y 14 semanas).

Pero no todo son vacaciones. Frente a este recorrido por las fiestas que tiene cada país, nos parece significativo valorar los días lectivos que se le atribuyen a cada uno de ellos; y ciertamente, para analizar los índices de calidad de la enseñanza o del sistema educativo de un país, este elemento viene a valorar el nivel del mismo. Podemos encontrar en este sentido, datos tan significativos como los que corresponden a los países de Finlandia y Dinamarca, frente a España. Según datos de la OECD (2003), la enseñanza primaria mantiene en Finlandia un total de 190 días lectivos, también Dinamarca cubre entre 190 y 200 días de enseñanza al año, mientras que en España el número desciende hasta 180 días. Es curioso comprobar sin embargo, que estas cifras no son directamente proporcionales al número de horas de enseñanza por año atribuidas a cada país. Así, en general, los países que presentan un mayor número de días lectivos son sin embargo, los que tienen un menor número de horas de enseñanza por curso. En el siguiente cuadro podemos ver más específicamente cada uno de estos datos:

TABLA II. Número de horas de enseñanza por años

Número de horas de enseñanza por años	
	Educación Primaria
Alemania	784
Dinamarca	640
España	880
Finlandia	656
Francia	905
Italia	748
Portugal	772
Media de los países	807

Y por otra parte, los países con una mejor evaluación en la escala del Informe PISA 2003, son los que tienen menor número de horas de enseñanza por año. Parece más significativo para lograr una evaluación positiva, el llegar a contar con un mayor número de días lectivos en el curso, que un mayor número de horas al año. En esta misma línea, el Informe de la OECD de 2005, muestra el número de horas lectivas que tiene un alumno

desde los 7 a los 15 años, y podemos observar que en España es de 7.731, mientras que en Finlandia es tan solo de 6.126; es decir, los estudiantes del país que encabeza los resultados del Informe PISA 2003 tienen, sin embargo, unas 1.600 horas menos de clase.

Valorando nuevamente las vacaciones como una de las motivaciones de los aspirantes a maestro, no podemos por menos que considerar también, el número de horas trabajadas por el profesor a la semana. Vacaciones versus número de horas, representan el contraste que envuelve a esta profesión y que marcan buena parte de la realidad de la misma. Esta circunstancia difiere también de un país a otro, aunque destaca Alemania cuyo intervalo oscila entre 38,5 y 40 horas convenidas. En España y Dinamarca es algo inferior: 37,5 y 37, respectivamente.

Y a la vista de todo esto, las vacaciones son ¿mito o realidad?

## La duración de la carrera

¿Son realmente atractivos estos estudios por tratarse de una carrera corta?

El Magisterio es en nuestro país -y hasta este momento- una carrera de ciclo corto. Esta circunstancia ha permitido que sus egresados logren el acceso al mundo laboral con una mayor inmediatez. Posiblemente, y así se ha demostrado en diversas investigaciones (De la Rosa Acosta, 1969; Ortega y Varela, 1985; García Ortiz, 1983; Sánchez Lissen, 1999), se trata de un factor que ha favorecido la elección de esta carrera a muchos candidatos; sobre todo, en aquellos casos en los que imperaba el deseo o la necesidad de lograr un salario. Entre los datos que nos aportan los estudios anteriormente citados, la opción señalada se menciona en distinta proporción:

- La investigación de Bernardo de la Rosa (1969) con una muestra de 1.355 estudiantes concluyó que la principal motivación de elección de la carrera era para los hombres y las mujeres, el tratarse de una «carrera corta».
- En los años ochenta, García Ortiz (1983) obtuvo resultados similares en relación a la anterior investigación. Participaron 250 alumnos y concluyeron que los alumnos habían elegido en primer lugar estos estudios por tratarse de una «carrera corta», mientras que las alumnas situaban esta opción en segundo lugar entre las 10 alternativas posibles.
- En la investigación de Ortega y Varela (1985) y con una muestra de 998 sujetos, la opción «por ser una carrera corta» fue elegida en tercer lugar con un porcentaje del 15,5%.

- En la década de los años noventa, González Sanmamed realizó una investigación para conocer los motivos y razones para estudiar magisterio. En ella participaron 650 estudiantes que adujeron entre los motivos extrínsecos: «la corta duración de la carrera», opción que consideraron la más valorada.

Tradicionalmente, la carrera de magisterio se ha convertido para muchos universitarios en una carrera trampolín; una carrera que permitía acceder a otros estudios superiores tras finalizar la misma y en muchos caos, ya trabajando.

Como sabemos, aunque caminamos hacia la convergencia del Espacio Europeo universitario, existe hasta este momento una gran disparidad en el número de años de formación en los países europeos. Presentamos a continuación, algunos de los más significativos:

TABLA III. Número de años en la formación inicial del profesorado

Países	Número de años de formación
Alemania	3,5 años: 7 semestres (Diplom.) 3+2 años (Máster) 4 años (Bachelor) + 1 año (Máster)
Dinamarca	4 años
España	3 años
Finlandia	3 años (Grado Educación Infantil) 4 / 5 años Educación Primaria
Francia	2 años 1000 -1200 h. teóricas + 500 h. prácticas
Irlanda	3 años + 1 año de especialización
Portugal	4 / 5 años

Fuente: Libro Blanco Título de Grado en Magisterio. Vol. I. Agencia Nacional de Evaluación de la Calidad y Acreditación, 2004.

Tal como se puede comprobar, no solo difieren los años de dedicación sino la contribución de horas teóricas y prácticas que tiene estipulado cada país. Esta circunstancia viene marcando la tendencia que recoge la mayoría de informes dedicados a la implantación del Espacio Europeo en la Universidad.

Pues bien, ¿me inclino a elegir el magisterio porque se trata de una carrera corta?; ¿mito o realidad?

## Requisitos para el acceso a la carrera

Ciertamente, no existe en nuestro país, grandes requisitos para acceder a la carrera de magisterio y esta circunstancia ofrece «vía libre» en muchos casos, para poder elegirla

con más facilidad que otras. Aunque se impone como requisito mínimo una nota de corte, sin embargo, muchos pensamos que esto no es suficiente.

Para algunos, puede ser un error no imponer algún requerimiento básico que logre captar –también– a los mejores estudiantes para esta carrera. Sin embargo, no parece fácil encontrar la mejor fórmula para lograr un reclutamiento justo y a veces, el único método radica en la realización de pruebas básicas y comunes que evalúan su selección.

Si el acceso es restringido, está controlado y valora los requisitos mínimos que debe tener un futuro docente, la elección de los candidatos podría ser más acertada que si se trata de un acceso totalmente abierto. Aunque este posicionamiento es legítimo, sin embargo, no existe unanimidad en esta afirmación, dado que se podrían quedar «en el camino» aquellos menos valiosos a priori, pero más válidos tras incorporarse y «beber» de la formación inicial.

J. M. Esteve (2003) reconoce que en nuestro país, la selección del profesorado no debería conformarse con unas oposiciones, sino que se debería valorar, además, las destrezas y la personalidad del candidato. En este sentido se expresa diciendo que:

La presencia de mayores tasas de ansiedad entre los profesores jóvenes y la aparición de una mayor incidencia de depresiones y neurosis, que se concentra en un grupo muy reducido, aconseja el empleo de mecanismos de selección de profesores basados en criterios de personalidad, tanto por el bien de los alumnos que puedan estar a su cargo como para evitar la situación de ruptura personal que sufren cada día muchos docentes en sus relaciones con los alumnos.

En otros países, los requisitos mínimos que se plantean en la selección y posterior acceso a la carrera, se contemplan de esta manera:

**TABLA IV.** Requisitos en el acceso a la formación inicial del profesorado

Países	Requisitos de acceso
Alemania	19 años, superar el examen Abitur
Dinamarca	Examen de acceso
Finlandia	Diploma educación secundaria + examen de acceso con entrevista
Francia	Tener Diploma Universitario de al menos 3 años de duración
Irlanda	Leaving Certificate Examination Algunos centros tienen entrevista individual
Luxemburgo	Diploma de Ed. Secundaria + examen de acceso. Tener menos de 35 años
Portugal	Diploma de Educación Secundaria

Fuente: OECD, 2003.

Al igual que alguno de los países anteriormente citados existen otros que también cuentan con examen de acceso; es el caso de Austria cuya prueba de acceso es el Reifeprüfung, o Polonia que tras obtener el certificado de enseñanza secundaria (Matura), tienen un examen de ingreso y una entrevista sobre actitudes y aptitudes, además de un certificado de salud.

En nuestro país, la selección se ofrece a posteriori ya que se realiza en el momento de acceso a la profesión docente. Como sabemos, esta circunstancia varía dependiendo del tipo de centro al que desee acceder. Si se trata de un centro privado o concertado, es competencia del director o en quién delegue, para elegir el profesor que deberá ocupar la plaza. Si se trata de un centro público, el acceso a la docencia se realiza a través de un concurso-oposición. Sin embargo, lo característico en este sentido es que no todos los países llevan a cabo dicho proceso selectivo en el mismo momento. En general, lo más habitual es que a este proceso se puedan presentar los estudiantes una vez que hayan finalizado la formación inicial; sin embargo existen algunos casos paradigmáticos como lo son Francia y Luxemburgo, en los cuales, el proceso de concurso público comienza antes de finalizar completamente el periodo de formación inicial. En ambos casos se trata de unas pruebas a las que concurren los candidatos «en el lugar de trabajo» antes de la fase final de cualificación.

En Francia, una de las selecciones más importantes tiene lugar antes de acceder al segundo curso en un IUFM. Concretamente, los estudiantes deben superar en primer lugar el Concours (CAPES o *Agrégation*) y posteriormente, el Comité responsable se encargará de evaluar a los candidatos y de realizar la selección de los mismos. A dicho Comité se le confía la redacción de un informe que incluye a todos los que han superado las pruebas y pasan a ser funcionarios (*Titulaire*).

TABLA V. Requisitos en el acceso a la profesión docente

Paises	Concurso-oposición	Pruebas de aptitud	Contrato temporal (si se supera un primer examen)
Alemania		X	
Bélgica	X		X
España	X		
Francia	X		
Grecia	X		
Luxemburgo		X	
Portugal			X

Fuente: Eurydice, 2004.

¿Son tan nimios los requisitos de acceso que se requieren en esta carrera, como para considerar este factor: un mito o una realidad?

## La vocación

Hasta ahora, los cuatro factores citados tienen un carácter extrínseco puro y temo, que el lector pueda pensar que hoy por hoy, aquéllos que desean ser maestros sólo valoran aspectos como el sueldo, las vacaciones, el número de años o la facilidad de la carrera. Ser maestros es algo más. En la elección se vuelcan otros motivos intrínsecos, motivos personales que nacen del interior y que, ciertamente, no sería justo olvidarlos en este artículo. Pues bien, entre ellos está la vocación.

Y la vocación es una llamada, un sentimiento, una tendencia o, sencillamente un servicio; en cualquier caso, todo aquello que se tache de verdaderamente vocacional, además de llenarse en lo personal, debe repercutir y procurar un valor social. Como señala Hansen (2001), «la vocación no es una devoción desinteresada» ya que en ella confluyen dos circunstancias: la realización de un servicio y la satisfacción en el cumplimiento de dicho servicio. En la fusión de ambas cosas radica el verdadero sentido vocacional de la enseñanza.

Aunque todo esto parezca indiscutible, sin embargo, ¿qué hay de vocación en la enseñanza?

En 2005, el informe de la OECD registró un descenso considerable en el número de profesores en toda Europa, al señalar: «que cada vez existen menos vocaciones, que la profesión envejece y que no somos capaces de generar nuevos candidatos».

Por otra parte, el estudio de la Comisión Europea: «Progresos hacia los objetivos de Lisboa en Educación y Formación», estima que será necesario reclutar en torno a 1.000.000 de nuevos profesores entre 2005 y 2015, dado el envejecimiento acelerado de esta población.

Parece tan evidente la disminución de profesores que se viene produciendo en muchos países europeos año tras año, que se han planteado interesantes Campañas de publicidad para captar nuevos docentes. Por ejemplo:

- En Francia, en el año 2001 se planteó el siguiente eslogan: «*Professeur, et si l'avenir c'étaitait vous?*: (Profesor, ¿y si tú fueses el futuro?)»<sup>7</sup>.
- En el Reino Unido (Inglaterra y Gales) en el año 2000 adoptaron el siguiente lema: «*Those who can, teach*»: Aquéllos que podáis, enseñad” (Eurydice, 2004 a, p. 118).

---

<sup>7</sup> En 1988, Francia planteó también una campaña de revalorización de la profesión docente y utilizó para ello el lema: «El porvenir de unos es el porvenir de los otros». Éste y otros reconocimientos los mencionan los profesores Esteve, Franco y Vera (1995) en el libro *Los profesores ante el cambio social*.

- En Andalucía, la Consejería de Educación emitió en 1992 un anuncio publicitario para elogiar el trabajo docente y atraer a nuevos candidatos. Éste se organizó en torno al lema: «Enseñar, qué gran tarea».

Estos pequeños ejemplos son una muestra más de la importancia que tiene esta profesión, pero también, de las dificultades que genera su tarea. La vocación puede ser un buen indicador, nos puede ayudar a tomar el pulso de la carrera, pero seguramente no es un indicador definitivo. Junto a ella se conjugan un amplio número de variables, a veces insalvables, que condicionan tremendamente el acceso a la misma.

Si tradicionalmente se ha vinculado al magisterio con la vocación, sin embargo, se trata de un elemento controvertible dado que no todos acceden a esta carrera llevados por la vocación. La vocación se asocia a los elementos más personales y se vincula, tal como señala Castaño (1983), a las connotaciones psicológicas como la motivación y autorrealización, u otras como la ansiedad y la información. Se trata también de una tarea cargada de deseos personales pero a la vez, de variables sociales y culturales que inciden en la última elección. Rivas (1988) considera que en esta decisión vocacional de la que hablamos, influyen aspectos como el realismo, la flexibilidad y el libre compromiso.

Desde nuestro punto de vista, el proceso vocacional de la docencia conjuga tres grandes momentos (Sánchez Lissen, 2003): uno *pre-vocacional*, otro *peri-vocacional* y finalmente, otro *vocacional* propiamente dicho. Los tres suelen estar presente en la vida de cualquier persona que se acerca a esta profesión, aunque para algunos, sin embargo, ha podido estar de manera latente. El primer momento, el *pre-vocacional*, nos sugiere una etapa de descubrimiento, de análisis y de acercamiento a algunas de sus funciones previas a la entrada en la universidad. Este momento ha sido calificado por Guerrero Serón (1993) como un periodo de «vocación temprana», incluyendo en el mismo a las clases particulares y a la experiencia que junto a ésta se desarrolla. En realidad, se trata de una actividad con varias interpretaciones ya que unos tratan de confirmar con este «entrenamiento» su propio entusiasmo por la enseñanza y otros quieren obtener, principalmente a través de ella, algún beneficio económico. Estas dos situaciones, muy lícitas por otra parte, pueden llegar a confundir al futuro enseñante sobre su verdadera función, sus tareas y la importancia que tiene el ejercicio de esta profesión. No obstante, esta facilidad para ejercer dicha actividad –sin ser maestro– ha devaluado tremendamente la labor docente. Sin embargo, también ha podido entusiasmar a otros y descubrir en él un oficio atractivo. El segundo momento, el *peri-vocacional*, surge o se afianza a lo largo de la formación inicial, es decir en las Escuelas de Magisterio, las cuales también tienen en su haber, el fortalecimiento del sentido vocacional del sujeto. Se trata de un momento de aná-

lisis y de comprensión de las tareas docentes y de su identificación con las mismas. Finalmente, cuando el diplomado accede al ejercicio de su profesión, y lo incorpora a su vida diaria, es cuando más explícitamente está manifestando un sentimiento *vocacional* hacia la docencia. Por ello, sería más idóneo hablar de «interés profesional», de «ilusión profesional» o incluso de «aspiración», en vez de vocación. Como podemos comprobar, entre el primer momento y el último, además de tiempo hay también maduración por parte de la persona, así como un mayor acercamiento a la carrera y posteriormente a la profesión; circunstancias todas ellas que llegan a enfatizar o a disuadir el interés por la misma.

Durante este proceso de formación podemos vislumbrar tres modelos de vocación que Ferrero (1994) denomina:

- *Vocación subjetiva*: si el sujeto siente atractivo por una actividad, pero carece de cualidades para ella.
- *Vocación objetiva*: posee esas cualidades, pero no el atractivo hacia ella.
- *Vocación auténtica*: si además de sentir atractivo por la actividad, tiene las cualidades suficientes para ello.

Ciertamente, lo ideal sería la vocación auténtica, lo que supondría que la persona logra sintonizar lo atractivo de la actividad con la formación de aptitudes personales hacia la misma. Pues bien, en este caso, debemos ayudar al alumno a despertar sus cualidades, a optimizar sus posibilidades y además, insistir en los aprendizajes que son necesarios para el desarrollo personal y profesional del futuro docente.

Pues bien, reiteramos una vez más la pregunta que venimos realizando en torno a cada uno de los factores que influyen en la elección de la carrera docente: ¿puede ser la vocación un mito, o es una realidad?

## Reflexión final: seguimos caminando

Impregnados del espíritu de Lisboa, el Consejo Europeo presentó en marzo de 2002 en Barcelona, los siguientes objetivos estratégicos:

- Mejorar la calidad y la eficacia de los sistemas de educación.
- Facilitar el acceso de todos a los sistemas de educación y formación.
- Abrir los sistemas de educación y formación al mundo exterior.

De una manera u otra, cada uno de estos objetivos están relacionados con el docente, con su tarea y su formación.

Hay un deseo muy claro y que está permanentemente aflorando en todos los discursos y es, hacer de ésta una profesión atractiva y procurar que lleguen a la enseñanza los mejores. Ya hemos advertido que uno de los problemas que viene rodeando en estos momentos a la enseñanza, es la jubilación de los profesores ya que contamos con una población adulta que en breve, se retirará de la profesión. Son casos muy significativos los de Alemania, Italia y Suecia ya que algo más de un tercio de docentes supera los 50 años de edad en estos países.

Con todo ello cabría resumir en tres acciones, las propuestas en torno a esta profesión.

### **Convertir la carrera docente en una profesión atractiva**

Si el número de estudiantes -en general amplio- que acceden cada año a las Escuelas de Magisterio, justifican el atractivo de una profesión, podríamos concluir que la profesión docente sigue atrayendo a un buen número de jóvenes. Este atractivo subyace en la propia carrera, además de generarle otras motivaciones que la profesión lleva aparejada. Desde este punto de vista cabría señalar tres momentos que pueden contribuir al interés por la misma; el primero está relacionado con la experiencia personal vivida durante su etapa de estudiante. Esa imagen seduce a algunos y desilusiona a otros; su propia experiencia alimenta la nostalgia e invita a muchos jóvenes a interesarse y acceder a la misma. El segundo momento corresponde a los años de estudio de la carrera, que debe ser una etapa lo suficientemente significativa como para despertar el gusto por el ejercicio de la profesión. En estos momentos en los que se pone en marcha el Espacio Europeo de Enseñanza Superior, la carrera de maestro debe poner las bases para lograr una formación sólida, bien equilibrada, que reporte las competencias necesarias y que sea de interés para los nuevos jóvenes que se acercan a ella. En este proceso se deben involucrar a los docentes en ejercicio, dado que su experiencia debe servir en la gestión de la nueva reforma y en su caso, reconducir el nuevo estilo profesional del colectivo docente. Y finalmente, el tercer momento se refiere a la práctica profesional. Tal como hemos comentado a lo largo del artículo, la enseñanza se encuentra rodeada de algunas situaciones incómodas o poco atractivas para el desempeño de la misma. En este sentido se debe arbitrar un pacto por la educación que proponga mejoras en las condiciones laborales de los docentes, o en aspectos relacionados con la promoción o la competitividad de los salarios.

Junto a ello, buena parte del atractivo o del desinterés de la profesión está muy relacionado con las salidas laborales de la carrera. En este sentido no son tan importantes los input como los output, dado que estos últimos son los que condicionan el interés por la profesión. Si hace más de 30 años el acceso a los estudios de magisterio vinculaban su atractivo con factores como: «tratarse de una carrera corta», «ser económicamente asequible» o «ser una carrera fácil» (De la Rosa Acosta, 1969), su elección en estos momentos está unida a factores relacionados con las posibles salidas laborales, con el sueldo o con la seguridad de futuro que ésta proporciona.

En cierto modo, todas estas motivaciones contribuyen a aumentar el atractivo por la profesión.

### **Procurar que lleguen a la enseñanza, también, los mejores estudiantes**

Esta segunda propuesta está estrechamente vinculada con la anterior. La profesión docente se merece tener buenos estudiantes, jóvenes brillantes que se interesen por el desempeño de esta función en la sociedad. Por el propio prestigio de la profesión y de todos los que se dedican a ella, debería quedar en el olvido aquella idea de que los malos estudiantes encuentran en el magisterio su salida profesional preferente.

Ciertamente es una idea preocupante, porque de ser así, estaríamos considerando la carrera de segunda categoría y a la profesión, de falta de calidad. Entre los indicadores de calidad, la formación e idoneidad del docente tiene un enorme valor y por tanto, asegurar la presencia de buenos estudiantes garantizará buena parte del éxito. Sin ser una relación directa, ya que existen otros elementos que también influyen en los niveles de calidad, sin embargo, sí es significativa y merece la pena tenerla en cuenta. Gracias a su valía debe obtener recompensas para el ejercicio de esta profesión y además, el mercado laboral le debe brindar determinadas oportunidades en ese ámbito.

En definitiva, para reclutar buenos alumnos, tanto en la carrera como después en la profesión, se le debe brindar al personal, ciertas mejoras que favorezcan a medio y largo plazo el desempeño de su trabajo.

La mayoría de países de la OCDE cuentan con dos modalidades de acceso a la profesión; una, el caso español, que consiste en acceder por *concurso-oposición* y la segunda, la modalidad de *convocatoria abierta*, que experimenta la mayor parte de ellos y entre los que se encuentran países como: Inglaterra, Irlanda, los países nórdicos, etc. Esta segunda opción tal como reconoce Francesc Pedró (2006), «permite gestionar mucho mejor los flujos de profesores a escala local y de centro».

## Mejorar la valoración social del profesorado

Los buenos resultados obtenidos por Finlandia en los Informes PISA de los años 2000 y 2003, han llevado a los expertos a estudiar cuáles son los elementos que justifican esos excelentes datos. En este sentido, convergen en reconocer que buena parte de su secreto está en la popularidad bien entendida de la carrera docente y de manera especial, en el prestigio social del maestro. En nuestro país, aún queda bastante por hacer en este sentido. Para lograrlo hay varios factores claves, centrados en la calidad y que están relacionados con aspectos como: la formación, la «autoconfianza» (Serrano, Lera y Contreras, 2007), la equidad y la eficiencia. Los nuevos planes de estudio que impulsa el Espacio Europeo de Enseñanza Superior deben ser sugerentes en este sentido y deben arbitrar las estrategias necesarias para alcanzar un mayor éxito en la consideración de la carrera y de la profesión.

Otro factor de peso en la valoración social del profesorado es el salario. Aunque ya hemos realizado diversas observaciones al respecto a lo largo del artículo, cabe reconocer que, aunque los datos del Informe de la OCDE (2006) señala que el salario de los docentes españoles es superior a la media de la OCDE, sin embargo, al realizar un análisis de los datos atendiendo a otros indicadores (el número de horas de trabajo, tareas realizadas, gestiones atribuidas, o el PIB per capita, entre otros) se comprueban las diferencias que existen con otras profesiones de mayor prestigio social. Se trata de un valor que pueden mejorar tanto las instituciones, los políticos como los propios profesionales. Confío plenamente en ellos para que a medio plazo, la profesión docente alcance la valoración que se merece.

A la vista de todos estos datos podríamos concluir señalando que el entusiasmo, el gusto y el deseo de acceder a la enseñanza y llegar a ser maestro, aunque no es un motivo suficiente para lograr el éxito en la profesión, sin embargo sí es un motivo necesario para desarrollar correctamente ésta -y posiblemente cualquier otra- tarea profesional.

## Referencias bibliográficas

- ANECA (2004). *Libro Blanco del Título de Grado en Magisterio*. Vol. 1. Madrid: Agencia Nacional de Evaluación de la Calidad y Acreditación.
- AZNAR MARTÍNEZ, A. (2005). Homenaje al maestro. Jornada de Homenaje al Maestro del Consejo Escolar de la Región de Murcia (103-108). Murcia: Consejería de Educación y Cultura, Consejo Escolar de la Región de Murcia.

- CAMPUZANO, A. (2005). Europa necesita profesores... pero faltan candidatos. *Revista del Colegio Oficial de Doctores y Licenciados en Filosofía y Letras y en Ciencias*, 167, 2-3.
- CASTAÑO LÓPEZ MESAS, C. (1983). *Psicología y Orientación Vocacional*. Madrid: Marova.
- DE LA ROSA ACOSTA, B. (1969). *Influencia de factores socioeconómicos en la elección de la carrera de magisterio*. Universidad de Valencia, Tesis doctoral.
- DELORS, J. (1996). *La educación encierra un tesoro*. Madrid: Santillana.
- ESTEVE, J. M. (1994). *El malestar docente*. Barcelona: Paidós.
- (2003). *La tercera revolución educativa*. Barcelona: Paidós.
- ESTEVE, J. M., FRANCO, S. Y VERA, J. (1995). *Los profesores ante el cambio social*. Barcelona: Anthropos.
- EURYDICE (2004 a). *La profesión docente en Europa*. Vol. 3. Informe II. Bruselas: Eurydice.
- (2004 b). *La profesión docente en Europa*. Vol. 3. Informe IV. Bruselas: Eurydice.
- FERNÁNDEZ ENGUITA, M. (2005). Profesores y centros: el individuo y la organización (133-137). XIX Semana Monográfica de la Educación. Educación de calidad para todos: iniciativas Iberoamericanas. Madrid: Fundación Santillana.
- FERRERO BLANCO, J. J. (1994). *Teoría de la educación: lecciones y lecturas*. Bilbao: Universidad de Deusto.
- GARCÍA ORTÍZ, R. (1983). *La elección de la carrera de profesor de EGB. Importancia del factor socioeconómico-familiar*. Universidad de Sevilla, Tesis Doctoral.
- GERVILLA CASTILLO, E. (1988). Educar Hoy: Profesión contra vocación. *Bordón*, 50 (1), 83-91.
- GONZÁLEZ SANMAMED, M. (1995). Motivos para estudiar magisterio: entre el idealismo y la profesionalización. *Adaxe*, 11, 65-75.
- GUERRERO SERÓN, A. (1993). *Maestras y maestros. Autonomía, práctica docente y sindicación en una profesión subordinada: un estudio sociológico*. Madrid: Ediciones Universidad Complutense de Madrid.
- HANSEN, D.T. (2001). *Llamados a enseñar*. Barcelona: Idea Books, S.A.
- LAGARES PÉREZ, S. (2000). *Factores influyentes en la elección de la carrera de Magisterio en Huelva*. Universidad de Huelva, Tesis Doctoral.
- MELGAREJO DRAPER, J. (2006). La selección y formación del profesorado: clave para comprender el excelente nivel de competencia lectora de los alumnos finlandeses. *Revista de Educación (Madrid)*, nº extraordinario, marzo, 237-262.
- ORTEGA, F. Y VARELA, J. (1985). *El aprendizaje de maestro*. Madrid: Servicio de Publicaciones del MEC.

- OCDE (2004). *Informe PISA 2003*. Madrid: Santillana.
- (2003). *Education Policy Analysis*. Paris: OECD.
- (2005). *Education at a glance 2005*. Paris: OECD.
- (2006). *Education at a glance: OECD Indicators-2006*. Paris: OECD.
- PEDRÓ, F. (2004). *Fauna académica. La profesión docente en las universidades europeas*. Barcelona: Editorial UOC.
- (2006). Un diagnóstico de la situación del profesorado en España desde una perspectiva comparativa. *Revista de Educación. (Madrid)*, 340, 243-264
- RIVAS, F. (1988). Asesoramiento vocacional: estado de la cuestión en las relaciones entre la teoría, la investigación y la aplicación. *Revista de Educación*, 286, 221-241.
- SÁNCHEZ LISSEN, E. (1999). *Factores condicionantes en la elección de la carrera de Magisterio*. Universidad de Sevilla, Tesis Doctoral.
- (2002). *Elegir enseñar: entre la vocación y la osadía*. Sevilla: Edición digital @tres. SL.
- (2003). La vocación entre los aspirantes a maestro. *Educación XXI*, 6, 203-222.
- SERRANO, J.A., LERA, A. Y CONTRERAS, O. (2007). Maestros generalistas versus especialistas. Claves y discrepancias en la reforma de la formación inicial de los maestros de primaria. *Revista de Educación (Madrid)*, 344, 533-555.
- SOCKETT, H. (1993). Education and will: Aspects of personal capability. *American Journal of Education*, 96, 195-214.
- VEGA GIL, L. Y OTROS. (2007). *Convergencias y divergencias de la formación inicial práctica del profesorado en la Europa Mediterránea*. Granada: Grupo Editorial Universitario.

## Fuentes electrónicas

- FINISH NATIONAL BOARD OF EDUCATION (2005). Quantitative educational Indicators. International comparisons of some features of Finnish education and training. Consultado el 4 de mayo de 2006, de [http://www.oph.fi/info/tilastot/International\\_comparisons.pdf](http://www.oph.fi/info/tilastot/International_comparisons.pdf)
- RIOBÓ, E. (2003). *La falta de profesores ya preocupa en la OCDE*. *Aprendemas.com*: Consultado el 11 de febrero de 2004, de [http://www.aprendemas.com/Noticias/html/N284\\_F03102003.HTML](http://www.aprendemas.com/Noticias/html/N284_F03102003.HTML).
- UNESCO (2004). *Educación para todos. El imperativo de la calidad. Resumen. Informe de Seguimiento de la EPT en el mundo*. París: UNESCO. Consultado el 4 de mayo de 2006, de <http://unesdoc.unesco.org/images/0013/001373/137334s.pdf> .

— (2006). *Los docentes y la calidad de la educación: seguimiento de las necesidades mundiales para 2015*. París, UNESCOPRESS. Consultado el 7 de enero de 2007, de <http://www.unesco.cl/esp/sprensa/noticias/194.act?menu=/esp/sprensa/>

**Dirección de contacto:** Encarnación Sánchez Lissen. Universidad de Sevilla. Departamento de Teoría e Historia de la Educación y Pedagogía. C/ Camilo José Cela, s/n, 41018. Sevilla, España. E-mail: [eslissen@us.es](mailto:eslissen@us.es)

# La pedagogía no tiene la culpa: un análisis de los problemas de la educación en España

## It's not pedagogy's fault: a paradigm switch to analyze the Spanish education system problems

Ana Benito Martín

Port Aransas City Hall, Texas, Estados Unidos

### Resumen

El descontento del profesorado respecto al sistema educativo es anterior a la reforma de 1990. Sin embargo, este descontento parece haberse recrudecido y materializado en una crítica y rechazo directo hacia la pedagogía constructivista. La reciente publicación del *Panfleto Antipedagógico* ha reunido eficientemente muchas de las quejas de un sector del profesorado en materia de educación y parece estar recibiendo gran apoyo de la comunidad educativa. Ni el contenido ni el tono del panfleto son novedosos, recoge una serie de lugares comunes presentes en la comunidad educativa como: «No se puede aprender de memoria», «los alumnos no tienen que esforzarse para aprender» o «igualar por abajo» que son inadecuadamente atribuidos a principios de la pedagogía constructivista. Aunque incorrectos, estos son asumidos por algunos miembros de la comunidad educativa como profesores, inspección educativa, pedagogos, asesores educativos, así como políticos y académicos en nuestro país. Además, la escuela comprensiva recibe muchas críticas por ser un sistema que tiene como efecto previsible e inherente un aumento de los problemas de disciplina y un descenso del nivel académico de los alumnos. Sin embargo, este modelo educativo ha sido implementado con éxito en otros países y, por tanto, las dificultades encontradas en su implantación son probablemente debidas al contexto cultural y socioeconómico español. El objetivo de este ensayo es redirigir las críticas a la pedagogía hacia problemas de la educación más tangibles y en ocasiones sugerir posibles soluciones. Pretende también reflexionar sobre algunas características de la sociedad

española y de su sistema educativo que permitan poner en contexto algunos de los problemas que muestra el alumnado. Entre ellos podemos mencionar: una fuerte orientación hacia modelos educativos de transmisión, escasa formación pedagógica del profesorado, bajos niveles de escolarización en el pasado, insuficiente inversión o una corta historia democrática.

*Palabras clave:* críticas a la pedagogía, malestar docente, reforma educativa, actitudes en educación, principios educativos, escuela comprensiva, memoria, motivación, formación del profesorado

### **Abstract**

Discontentment among Spanish teachers regarding the present education system predates the 1990 reform. However, this dissatisfaction seems to have greatly intensified manifesting itself in criticism against, and point-blank rejection of, constructivist pedagogy. The recent publication of the 'Antipedagogical Pamphlet' (*Panfleto Antipedagógico*) efficiently assembled in one document many of the complaints of one sector of Spanish teachers regarding education matters and it appears to be receiving widespread support from the educational community. Neither the content nor the tone of the pamphlet are novel. Rather, it gathers a number of common points of view among teachers, such as: 'It is not possible to learn by memorizing facts', 'students should not force themselves too hard to learn', or 'standards should not be equalised down to a lowest common denominator', all of which are inaccurately attributed to the principles of constructivist pedagogy. In spite of being untrue, these myths are accepted by some members of the educational community such as teachers, educational experts, policy makers, politicians and academics in Spain. Furthermore, comprehensive schools are severely criticised as they are understood as an educational system for which the lack of discipline and a perceived plummeting of students' academic achievement constitutes a foreseeable, inherent and inevitable effect. However, this educational model has proved to be successful in other countries; consequently, the difficulties experienced in its implementation in Spain might probably be due to the cultural and socioeconomic context. This essay intends to redirect unfavourable judgements to pedagogy towards more tangible problems and, occasionally, to suggest possible solutions. It also intends to reflect upon some features of Spanish society and its education system to place in context some of the problems observed in Spanish students, i.e.: a strong orientation toward transmission modes of delivery, a scarce pedagogical background of Spanish teachers, low schooling levels in the past, insufficient investment in education, or a short history as a democratic country.

*Key Words:* criticism to pedagogy, teachers' discontentment, educational reform, educational attitudes, educational principles, comprehensive school, memory, motivation, teacher's training

## Introducción

El «malestar docente» parece haberse convertido en un movimiento de rechazo a la Pedagogía. Este fenómeno está representado con eficiencia en el *Panfleto Antipedagógico*, que está recibiendo apoyo de un sector de profesores e intelectuales. Los reiterados comentarios en los medios de comunicación en contra de la pedagogía y de la situación de la educación, parece que han dado lugar a un movimiento en contra de la pedagogía caracterizado por:

- Un rechazo a ciertos «principios pedagógicos» inexactos, pero que parecen haberse asumido a fuerza de repetirlos como «no se puede aprender de memoria»; «los alumnos no tienen que esforzarse»; «hay que rebajar el nivel» o «igualar por abajo».
- Rechazo a la escuela comprensiva como un sistema de garantía de fracaso.
- Creencia en la necesidad de restaurar la autoridad en el aula.
- Menosprecio a la formación pedagógica.

Este ensayo trata de aclarar como los *Mitos de la Logse* (Benavente Barreda, 2001) son más bien «leyendas urbanas» que realidades de la pedagogía. Además, trata de esbozar posibles razones del «fracaso» de la escuela comprensiva en España. Puesto que la escuela comprensiva hasta los 16 años en otros países, en especial en Finlandia, ha sido aplicada con excelentes resultados (Finish PISA Team, 2002), el fracaso no se debe atribuir al sistema comprensivo en sí mismo, sino a otros factores entre los que se pueden incluir: factores sociopolíticos y económicos, corta tradición democrática, insuficiente inversión en educación o una deficiente formación pedagógica del estamento educativo.

## Contenidos y memoria

Uno de los más importantes malentendidos es la interpretación de que el alumno no debe aprender de memoria. Benavente Barreda (2001, p. 95) señala que «El principio básico del cognitivismo es desterrar la memoria y fomentar el aprendizaje racional». El cognitivismo surgió como paradigma contrario al del conductismo y se preocupa por

los procesos internos del aprendizaje, es decir, por cómo el cerebro aprende. Los constructivistas, especialmente aquellos que se encargan de estudiar cómo se produce el conocimiento en el individuo (Phillips, 1995), creen que para que los aprendizajes tengan sentido y puedan recuperarse, deben estar relacionados con aprendizajes previos, además se debe de disponer de una estructura mental organizada necesaria para hacer esos aprendizajes disponibles (Donovan, Bransford, Pellegrino, 1999; Winn, 2002). Para ellos el aprendizaje es un proceso de elaboración y no de acumulación de conocimiento (Proulx, 2006). Esto no significa que no se deba usar la memoria, sino que ésta no debe ser la única forma de aprender.

Es difícil saber cómo comenzó a diseminarse esta idea, pero la cita al principio de este capítulo, tomada de una revista educativa sometida a revisión por pares, invita a la reflexión. En ella se afirma de forma errónea, y sin citar referencias, que el objetivo del cognitivismo es desterrar la memoria. Esto evidencia ciertas deficiencias en la comprensión y transmisión de los principios pedagógicos ya al principio de la cadena educativa, las fuentes primarias de información.

Otra idea equivocada es considerar el constructivismo como una teoría que determina o recomienda cómo se enseña e incluso un método didáctico. Benavente Barreda (2001, p. 95) subraya que uno de los principios (sic) de la LOGSE es «la aplicación del cognitivismo constructivista, con las consecuencias didácticas que su adopción implica (...) influyendo en la metodología didáctica de cualquier área o asignatura, e incluso en el papel del profesor dentro del aula». El constructivismo puede ofrecer implicaciones pedagógicas, pero no proporciona orientaciones didácticas porque el constructivismo es una teoría de cómo se aprende y no de cómo se enseña. Orienta más en qué no hacer que en qué hacer en el aula (Proulx, 2006; Winn, 1996). En resumen, los constructivistas consideran que el hecho de presentar una serie de hechos bien estructurados no va a promover necesariamente el aprendizaje. Esto no implica que se deba eliminar este método didáctico, sino que se pueden usar otros.

La razón de esta confusión puede estar relacionada con la falta de rigor pedagógico en los materiales destinados a la formación y apoyo de profesores. En un estudio (en lengua inglesa) se encontró que los materiales eran muchos y contenían ideas sobre «la enseñanza constructivista» que poco tienen que ver con el marco teórico (Rubie-Davis, 2006). Ni siquiera este es el caso en nuestro país donde las propuestas metodológicas fueron escasas y muy generales y su desarrollo posterior, en materiales didácticos, páginas Web, libros etc., ha sido escaso en especial en las primeras etapas de la reforma y su difusión no generalizada.

## Disfrutar y esforzarse no son antónimos

La «literatura antipedagógica» ha extendido la idea de que con la «nueva Pedagogía» los alumnos no tienen que esforzarse. El *Informe Pisa 2003* dedicó un capítulo a evaluar el aprendizaje de las Matemáticas en relación con las actitudes y estrategias cognitivas y metacognitivas de los alumnos. Los resultados de su estudio coinciden con los de otras investigaciones: «Sugieren que los estudiantes se embarcarán (...), en un aprendizaje de alta calidad y usando estrategias variadas si están motivados, no están ansiosos en relación al aprendizaje y creen en sus propias capacidades» (Organization for Economic Co-operation and Development, 2004, p. 156). Este aprendizaje de alta calidad requiere esfuerzo por parte del alumno que solo hará si se siente motivado y disfruta aprendiendo. Por lo tanto, el objetivo de disfrutar aprendiendo es conseguir que el alumno se esfuerce.

El concepto de motivación es otro de los malentendidos. La Pedagogía denomina algunos conceptos con palabras del lenguaje cotidiano (Davis y Sumara, 2003) y la falta de comprensión profunda de estos conceptos puede dar lugar a malas interpretaciones. Por ejemplo, Moreno Chumillas (2006) señala:

La voluntad no cotiza en el ámbito escolar, se ha transmutado en motivación, (...) la primera es disciplina de pupitre (...) y alude a una actitud personal tendente al propio esfuerzo (*a Dios rogando...*), mientras la segunda aparece como factor exógeno cuya carencia se achaca al incompetente profesor (para. 14).

Sin embargo, la motivación que se busca desde el punto de vista educativo es una motivación intrínseca. Ésta es un estado interno que estimula al aprendizaje y lo dirige. Consiste en competencias tales como la persistencia, establecer objetivos y retrasar la gratificación (Gregory y Chapman, 2003, p. 7). Un profesor no puede transmitir estas competencias a sus alumnos, pero sí puede intentar crear el ambiente adecuado y las actividades didácticas apropiadas para promoverlas.

Durante cuarenta años los investigadores han encontrado que las expectativas de los profesores sobre un alumno en particular, afectan a su aprendizaje (Rubie-Davies, 2006). La misma autora encuentra que las expectativas de los profesores sobre el grupo en conjunto afectan los resultados del conjunto de los estudiantes. La imagen que con frecuencia se da en los medios del alumnado y de la juventud española en la prensa resulta desoladora:

Alumnos carentes de principios y que le ponen el freno a valores tales como la amistad, la solidaridad, el compañerismo, el respeto a los otros, la aceptación

de los estudios, etcétera (...) muy mal educados, violentos, intolerantes, egoístas, ingratos, falsos, irresponsables, vagos... (Alberto Rodríguez citado por Fez Marrero, 2006, p. 8).

Comentarios como este están omnipresentes en la prensa. Carbonell y Tort, (2006, p. 21) señalan que «los medios privilegian (...), los acontecimientos excepcionales que rompen la normalidad y continuidad cotidianas». Es posible que el sensacionalismo mediático esté colaborando a proporcionar una visión distorsionada de la realidad educativa y de sus actores. Porque aunque los resultados académicos de nuestros alumnos y alumnas están significativamente por debajo de la media de los países de la OECD, están distribuidos en todos los percentiles, esto es un poco de todo (OECD, 2004). Es probable que los resultados sean los que cabe esperar en un país que hasta hace poco tenía las tasas más altas de analfabetismo, las tasas más bajas de escolarización y que invierte menos en educación que los países de su entorno (OECD, 2006), y con un sistema educativo con una fuerte orientación hacia la transmisión de contenidos, cuando el *Informe PISA* evalúa capacidades.

## La autoridad del profesor

Una parte importante de nuestra sociedad parece tener claro que el problema en las aulas es que se ha desposeído al profesorado de autoridad.

Era de esperar que un mayor número de alumnos escolarizados con diferentes necesidades educativas fuera a provocar problemas en las aulas y así ha sido desde los primeros años de implantación de la LOGSE (y antes). Se ha tardado 16 años en diseñar una estrategia nacional para afrontar el problema, aunque algunos institutos, comunidades autónomas y sindicatos llevan más tiempo analizando soluciones. Los materiales de la *Página de Convivencia* del Ministerio de Educación y Ciencia, que provienen de diversas fuentes, sindicatos, organizaciones, páginas personales, experiencias de institutos, etc., muestran un gran consenso en que las soluciones a los problemas de disciplina se encuentran en la creación de centros con un funcionamiento democrático, que propicien un clima de diálogo y que eduquen la responsabilidad de los alumnos. Entre las propuestas se encuentran la creación de comisiones de convivencia, mediación, ayuda entre iguales, negociación y diálogo, toma de decisiones, desarrollo de la empatía y de la

escucha activa, resolución de conflictos, metodologías participativas etc. Sin embargo, los sectores del profesorado más descontentos y más activos en sus críticas al sistema educativo parecen demandar modelos de subordinación del alumno frente a la autoridad del profesor, la posibilidad de aplicar castigos y la segregación de los alumnos con historia de indisciplina como ilustran los siguientes ejemplos:

Y los que no quieren, también estudiarían más para librarse de unos castigos que, si han de funcionar como tales, les tendrían que resultar más fastidiosos que el propio estudio. (Moreno Castillo, s. f., p. 21).

Sin perjuicio de la opción, por decisión judicial, del internamiento en correccionales (...) se crearán, para los alumnos de edad inferior a 14 años que persistentemente ocasionen conflictos en sus colegios e institutos, centros de reeducación con normativa específica y profesorado voluntario (Fez Marrero, 2006, p. 19).

Las propuestas del MEC requieren planificación de estrategias en el ámbito del centro, la colaboración de los profesores, el dominio de las técnicas y estrategias mencionadas y la convicción de que para la democracia se educa con medios democráticos. Es esperable que haya una fuerte oposición a estos planes, ya que «el sector crítico» piensa que esta línea «blanda» de pensamiento es la responsable de la actual situación:

Pero lo más sorprendente de la LOE es que cuando ya (...) se hace evidente la necesidad de recuperar el valor educativo de la disciplina, del esfuerzo o de la exigencia, los socialistas insisten en resolver los problemas con medidas inspiradas en lo que se ha dado en llamar «angelismo» escolar, causa principal de los males actuales de la escuela pública. Las soluciones que se proponen sólo se explican si están dictadas por el sectarismo y la demagogia (Delibes, 2005, p. 23).

## «Igualar por abajo»

Benavente Barreda (2001) afirma que «una de las más graves deficiencias de la enseñanza comprensiva es la de tener que rebajar los niveles de exigencia». También indica que la LOGSE venía matizada por la atención a la diversidad. Para ser exactos, la atención

a la diversidad era un principio de la LOGSE, no un matiz. La necesidad de atender a la diversidad se basa en la noción de que el aprendizaje se produce en la zona donde se desafíe intelectualmente al alumno sin producirle ansiedad, esa zona de desarrollo proximal según Vygostky (Smith, Cowie y Blades, 2002, p. 429) es diferente para cada alumno. Por lo tanto, para asegurar el aprendizaje de todos los alumnos y alumnas se proponía la creación de actividades didácticas con diferentes grados de dificultad, variadas para acomodar los diferentes estilos sensoriales, cognitivos y de aprendizaje de los alumnos (Ginnis, 2002; Gregory y Chapman, 2002). Se contemplaban las adaptaciones curriculares tanto para los alumnos con necesidades educativas especiales o que promocionaban con pendientes e incluso para los aventajados y superdotados. El «listón» no está más bajo, está situado donde los objetivos del profesor lo hayan marcado. La teoría pedagógica indica que hay que partir del nivel de cada alumno o alumna e intentar desarrollar al máximo sus capacidades, es decir, la propuesta es ajustar el nivel para que se adapte a cada alumno y no a todos los alumnos a un único nivel adaptado al más bajo. No obstante, los profesores reiteran sus quejas sobre las presiones que sufren para aprobar a más alumnos como señala Moreno Chumillas (2006, para. 6): «aumentar la presión social sobre el profesor hasta que claudique en la Junta de Evaluación».

De nuevo encontramos discrepancias entre las intenciones de la Pedagogía y las interpretaciones que de ellas se han hecho. Las medidas de atención a la diversidad implican una atención individualizada y no una rebaja general de la exigencia. Otro debate es si este modelo educativo es aplicable o realista en nuestro contexto educativo y social. Todos los que hemos estado en un aula con alumnos con problemas de comportamiento, con necesidades educativas especiales, repitiendo curso o con áreas pendientes, sabemos el desafío que supone atender a la diversidad. Son muchos los que argumentan que es imposible. Sin embargo, esto no afecta a la noción de que los alumnos aprenden mejor si se considera su punto de partida, sus estilos de aprendizaje y si participan de forma activa como señala la investigación. Es decir, no afecta al planteamiento pedagógico. Y es posible que no sea aplicable, pero no se puede asegurar mientras no se planteen, implementen y evalúen reformas que afecten a la práctica en el aula. En el contexto actual es difícil, pero es posible que algunas medidas permitieran aplicar un modelo comprensivo; entre otras: disminuir la ratio de alumnos, en especial en las clases conflictivas y con alumnos con necesidades educativas, introducir un profesor de apoyo en el aula, crear la figura de ayudantes de aula, crear aulas temáticas dotadas de materiales adecuados para la diversidad de alumnos y que permitan hacer agrupamientos flexibles o modificar horarios. En las sucesivas reformas se echa de menos la implementación y evaluación de medidas que permitan crear las condiciones en que este modelo podría ser aplicable.

## Comprensividad

Los detractores de la escuela comprensiva proponen la aparición de itinerarios para acomodar a los alumnos con distintas «capacidades». Los itinerarios son poco más que una vuelta a la separación entre lo profesional y lo académico. Una constante es que mencionan reiteradamente el fracaso y abandono de la escuela comprensiva en Gran Bretaña. Sin embargo, existe una gran diferencia entre la propuesta de cambio española y la británica. Mientras en España se pretende volver a un modelo anterior, la reforma Británica es un paso adelante (esta opinión se refiere sólo a los aspectos pedagógicos y no a la organización o modelos de financiación) y no una vuelta a los modelos del pasado. Se pasa de la atención individualizada a la educación a medida (*tailored*), sin cisma entre lo profesional y académico, consideran que todos los alumnos tienen capacidades, que se debe hacer un esfuerzo para mantenerlos en el sistema educativo y se apoyan las necesidades educativas especiales en toda la etapa (hasta los 19 años) (Department for Education and Skills, 2005).

Por otro lado, son escasos los comentarios sobre el por qué del éxito del sistema comprensivo finlandés en la última evaluación PISA (OECD, 2004). Resultados aparte, el sistema educativo finlandés es parecido al nuestro en cuanto a planteamientos y organización: 9 años (de los 7 a 16 años) de escolarización comprensiva, las necesidades educativas en su mayoría se apoyan en el aula, no hay repetición, no hay exámenes estandarizados, etc. (Ministry of Education, 2006). El éxito parece estar relacionado con las características de la sociedad finlandesa. Los alumnos parecen estar motivados, interesados y confían en sus propias capacidades y tienen gran afición a la lectura, es una sociedad homogénea y con elevados niveles de equidad, invierten mucho en educación en especial en educación primaria y sus profesores están altamente cualificados (Institute for Educational Research, 2002; Melgarejo Draper, 2006).

## La formación del profesorado

Dejando al margen las razones socioeconómicas, mientras en Finlandia se considera que para hacer frente a grupos heterogéneos un profesor debe ser «un experto altamente educado en pedagogía». En España, con frecuencia, se oyen voces que muestran desprecio hacia la formación pedagógica. Se argumenta que para dar clases sólo hace falta tener conocimientos de la materia. Según Ricardo Moreno (s. f.):

Enseñar se parece más a un arte que a una ciencia (...) el resto depende de la afición del profesor por el saber que se pretende transmitir, de la capacidad de ser claro y ordenado en la exposición, de la de hacerse respetar por los alumnos y comunicar con ellos. Para quien carece de estas habilidades los cursos de formación pedagógica son inútiles, para quien las tiene son superfluos.

Algunos intelectuales del país expresan, también, su rechazo a la formación pedagógica. Rodríguez Adrados (2006, para. 13) miembro de la Real Academia de la Lengua escribe respecto a la implantación del futuro máster necesario para ejercer la docencia: «Nosotros, los concededores del tema, somos suficientes» Sobre el mismo tema los decanos de filosofía «critican que esa formación se haga "al margen" de las facultades donde se forman los futuros licenciados (...) se quejan, además, de que a la formación disciplinar de este master sólo se destinarán 12 de los 60 créditos de que constará el curso» («Los decanos...», 2006, para. 2, 3).

Una parte del profesorado y de la sociedad española parece no haber superado el modelo tradicional de educación. No obstante, ese contexto no es en el que nos encontramos. El modelo educativo actual centrado en la actividad del alumno, con objetivos basados en capacidades, con un alumnado diverso, con problemas de comportamiento, etc. requiere, nos guste o no, una formación y habilidades distintas a las que requería un profesor en el pasado. Entre los nuevos desafíos con que se encuentran los profesores actuales están tratar con alumnos más variados y conflictivos, diseñar unidades didácticas y actividades de aprendizaje, participar en la elaboración de las programaciones de departamento, elaborar adaptaciones curriculares para los alumnos de necesidades educativas especiales, incorporar las nuevas tecnologías en el aula, etc. El profesor debe por tanto actualizarse en su materia en un mundo rápidamente cambiante, conocer técnicas de dinámicas de grupos y de manejo del aula y del comportamiento, tener conocimientos de desarrollo del currículo y de problemas de aprendizaje, y crear y seleccionar materiales. Todas las decisiones de un profesor deben estar basadas en qué es mejor desde el punto de vista del proceso de enseñanza-aprendizaje y se pueden tomar confiando en la intuición o aceptando la validez de las investigaciones pedagógicas.

La aplicación de la pedagogía puede ser difícil en el aula, incluso en las condiciones más favorables. Las reformas educativas en España, hacen y deshacen itinerarios, ponen y quitan asignaturas, reordenan las enseñanzas y hacen y deshacen listas de contenidos. Las miles de alegaciones a las reformas provenían de sectores que, como es lógico, defendían sus intereses, entre otros: la distribución de horas lectivas de diferentes asignaturas, la consideración de una asignatura como obligatoria, conciertos o

mejora de las condiciones salariales. Se echaron de menos, sin embargo, alegaciones que fueran al centro de los problemas que padecen los profesores en el aula y de los que un amplio sector del profesorado se queja.

Es probable que los profesores sigan, después de su implantación, sin tener instrumentos para enfrentarse a los problemas de comportamiento, a la diversidad, a los problemas de aprendizaje, a las necesidades educativas especiales, a las deficiencias del currículo y a la inadecuada infraestructura y recursos para implementar el modelo educativo que se les demanda.

Sin embargo los profesores, que son los que conocen y sufren estas dificultades, no van a ser interlocutores válidos mientras no estén preparados para hablar de educación con un conocimiento sólido de la teoría educativa, usando un discurso apropiado y no recurriendo a lo que Horn (2002) denominó conocimiento indígena (traducción del inglés *indigenous knowledge*), que es el que se adquiere como resultado de la experiencia.

No son los actuales paradigmas educativos los que están en tela de juicio y si lo estuvieran sería porque nuevas investigaciones arrojan nueva luz sobre la materia y permiten crear un nuevo paradigma. En nuestro país, como este trabajo muestra, la interpretación, transmisión y aplicación de la teoría educativa ha tenido algunas deficiencias. Sería conveniente realizar un análisis profundo, reflexivo y desinteresado de los problemas de la educación, así como, dialogar con los miembros de la comunidad educativa para acordar, evaluar e implementar soluciones. La habilidad de analizar y evaluar problemas, de pensar críticamente, de reflexionar son algunas de las capacidades que las nuevas teorías pedagógicas pretenden estimular en el alumnado. Es probable que las intenciones no sean suficientes cuando se carece de modelos adecuados.

## Referencias bibliográficas

- BENAVENTE BARREDA, J. M. (2001). Los mitos de la reforma. *Revista Iberoamericana de Educación*, 27, 77-98.
- CARBONELL, J. Y TORT, A. (2006). *La educación y su representación en los medios de comunicación*. Madrid: Ediciones Morata.
- DAVIS, B. & SUMARA, D. (2003). Why aren't they getting this? Working through the regressive myths of constructivist pedagogy. *Teaching Education*, 14, 123-140.
- GINNIS, P. (2002). *The teacher's toolkit. Raise classroom achievement with strategies for every learner*. Carmarthen: Crown House publishing.

- GREGORY, G. H. & CHAPMAN, C. (2002). *Differentiated Instructional Strategies. One Size Doesn't Fit All*. Georgia: Corwin Press.
- HORN, R.A.J. (ed.) (2002). *Understanding Educational Reform, a reference handbook*. Santa Barbara: ABC-CLIO, Inc.
- MELGAREJO DRAPER, J. (2006). La selección y formación del profesorado: clave para comprender el excelente nivel de competencia lectora de los alumnos finlandeses. *Revista de Educación*, Extraordinario, 237-262.
- ORGANIZATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT (2004). *Learning for tomorrow's world first results from PISA 2003. Executive Summary*. Paris: OECD publishing.
- ORGANIZATION FOR ECONOMIC COOPERATION AND DEVELOPMENT (2006). *Education at a glance, indicators*. Paris: OECD publishing.
- PHILLIPS, D.C. (1995). The good, the bad, and the ugly: The many faces of Constructivism. *Educational Researcher*, 24, 5-12.
- PROULX, J. (2006). Constructivism: a re-equilibration and clarification of the concepts and some potential implications form teaching and pedagogy. *Radical Pedagogy*, 8, 1.
- RUBIE-DAVIES, C. M. (2006). Teacher expectations and student self-perceptions: exploring relationships. *Psychology in the Schools*, 43, 537-552.
- SMITH, P.K., COWIE, H. & BLADES, M. (2002). *Understanding children's development. 3<sup>rd</sup> Edition*. Cornwall: Blackwell Publishers.
- WINN, W. D. (1996). *Cognitive Perspectives in Psychology*. En D. JONASSEN (ed.), *Handbook of Research on Educational Communications and Technology* (112-142). New York: Oxford University Press.

## Fuentes electrónicas:

- DELIBES, A. (2005). En nombre de la equidad no se puede extender la ignorancia. Papeles FAES, 21. Consultado el 10 de abril de 2006, de [http://www.fundacion-faes.es/boletin/boletin.cfm?id\\_seccion=1507](http://www.fundacion-faes.es/boletin/boletin.cfm?id_seccion=1507).
- DEPARTMENT FOR EDUCATION AND SKILLS. (2005). *14-19 Education and Skills*. London: DefES. Consultado el 10 de agosto de 2006, de <http://www.dfes.gov.uk/publications/14-19educationandskills/pdfs/14-19WhitePaper.pdf>.
- DONOVAN, M. S., BRANSFORD J. D. Y PELLEGRINO, J. W. (Eds.). (1999). *How people learn: bridging research and practice*. Washington, DC: National Academy of Sciences. Consultado el 10 de septiembre de 2006, de <http://www.nap.edu/catalog/9457.html>.

- FEZ MARRERO, DE E. (2006). Educación: «volver al BUP y COU», la única solución. *La Voz de Lanzarote*, 25. Consultado el 10 de agosto de 2006, de <http://www.lavozde-lanzarote.com/spip.php?article6537>.
- INSTITUTE FOR EDUCATIONAL RESEARCH (2002). *The success of Finnish in PISA, and some reasons behind them*. University of Jyväskylä: Finnish PISA Team. Consultado el 10 de Julio de 2007, de <http://ktl.jyu.fi/pisa/publication1.pdf>.
- MINISTRY OF EDUCATION (2006). *Education and Science in Finland*. Finland: Helsinki University Press. Consultado el 10 de septiembre de 2006, de [http://www.minedu.fi/export/sites/default/OPM/Julkaisut/2006/liitteet/eng\\_opm15.pdf?lang=en](http://www.minedu.fi/export/sites/default/OPM/Julkaisut/2006/liitteet/eng_opm15.pdf?lang=en).
- MORENO CASTILLO, R. (s.f.). *Panfleto antipedagógico*. Consultado el 10 de julio de 2007, de <http://www.mat.ub.es/~cerda/Pamflet.pdf>.
- MORENO CHUMILLAS, E. (2006). *La educación Zen*. Consultado el 10 de septiembre de 2007, de <http://amesweb.tripod.com/educacionzen.htm>.
- RODRÍGUEZ ADRADOS, F. (2006, 12 de junio). Ni griego ni latín. *El País*. Consultado el 10 de Julio de 2007, de [http://www.elpais.es/articulo/educacion/griego/latin/elpedupor/20060612elpepedu\\_9/Tes/](http://www.elpais.es/articulo/educacion/griego/latin/elpedupor/20060612elpepedu_9/Tes/).

Los decanos de Filosofía piden anular el 'máster' para ser profesor de secundaria.(2006, 12 junio). *El País*. Consultado el 10 de agosto de 2006, de [http://www.elpais.es/articulo/educacion/decanos/Filosofia/piden/anular/master/ser/profesor/s\\_cundaria/elpedupor/20060612elpepedu\\_5/Tes/](http://www.elpais.es/articulo/educacion/decanos/Filosofia/piden/anular/master/ser/profesor/s_cundaria/elpedupor/20060612elpepedu_5/Tes/).

**Dirección de contacto:** Ana Benito Martín. Port Aransas City Hall. Texas, EEUU. 230 S. Cut-Off Rd., Aransas Harbors, Apt. 121, Port Aransas, Texas 78373. USA. E-mail: [anabenitoes@yahoo.es](mailto:anabenitoes@yahoo.es)





## **Experiencias educativas (Innovación)**



# Participación y rendimiento del estudiante universitario en un proyecto docente interdisciplinar, bilingüe y virtual

## Participation and performance of university students in an interdisciplinary and bilingual e-learning project

Manuel Cuadrado

María Eugenia Ruiz Molina

*Universidad de Valencia. Facultad de Economía. Valencia, España*

Mercedes Coca

*London School of Economics. Londres, Inglaterra*

### Resumen

La enseñanza y el aprendizaje han empezado a aprovechar desde hace unos años los beneficios de la globalización y de las nuevas tecnologías, favoreciendo así la colaboración entre universidades de diferentes contextos. Los idiomas no son una excepción en este proceso. Prueba de ello resulta el presente estudio que plantea como objetivo analizar los resultados de un novedoso proyecto de enseñanza-aprendizaje interdisciplinario y bilingüe entre dos universidades europeas: la London School of Economics (Reino Unido) y la Universidad de Valencia (España). Un proyecto formulado con la idea de impulsar el conocimiento de un idioma extranjero y fomentar la cooperación y el intercambio de información entre los estudiantes de estas dos universidades a través de un entorno virtual. La misión de los estudiantes consistía en afianzar una segunda lengua así como sus conocimientos de economía y administración de empresas tanto en español como en inglés. A partir de los diversos datos de participación y las calificaciones finales de los estudiantes en sus respectivas materias, calculamos coeficientes de concordancia con el objetivo de analizar la participación y el rendimiento de éstos. Como resultado, se observa una relación positiva y significativa entre la implicación de los estudiantes en el proyecto y sus calificaciones finales para los estudiantes de ambas instituciones. Los

resultados demuestran que el proyecto contribuye de forma positiva en el proceso de enseñanza-aprendizaje tanto de idiomas como de otras disciplinas.

*Palabras clave:* recursos didácticos, Internet, docencia, idiomas, rendimiento del alumno, programa interuniversitario

### **Abstract**

Education has recently begun to take advantage of the benefits of the globalization and the application of new technologies, thus facilitating the collaboration between universities of different contexts. Language teaching is not an exception in this process. As a result of this, the present study aims at analyzing the results of an innovative interdisciplinary and bilingual e-learning project between two European universities: the London School of Economics (United Kingdom) and the University of Valencia (Spain). A project was formulated, with the purpose of fostering the knowledge of a foreign language and the cooperation and exchange of information between the students of these two universities through the use of a virtual platform (webCT). The students' remit consisted of strengthening their applied knowledge of a foreign language, as well as of economy and business administration, both in Spanish and English. From the diverse data on students' participation and final qualifications, we calculated rank correlation coefficients in order to analyze students' participation and performance. Study results show a positive and significant relation between students' involvement in the project and their final qualifications within both institutions. Thus, the project contributes in a positive way in the teaching-learning process of languages as well as other disciplines.

*Key Words:* didactic resources, Internet, teaching, languages, performance, interuniversity programme

## **Introducción**

Las universidades vienen colaborando entre sí desde hace ya años y a través de diversos programas. En este sentido, se han realizado esfuerzos continuos para desarrollar acuerdos de colaboración no sólo entre las instituciones educativas de la Unión Europea, sino también a nivel mundial. Estos acuerdos han facilitado el intercambio de miles de estudiantes y docentes. El establecimiento de redes ha permitido también la cooperación interuniversitaria. Entre otras, la Asociación de Universidades Europeas (EUA), la Comunidad de Universidades Mediterráneas (CUM) o

la Asociación Europea de Educación Internacional (EAIE) son una muestra de la voluntad de las universidades para trabajar con otras instituciones.

Un paso más en este proceso es la *Declaración de Bolonia*. Su finalidad ha sido homogeneizar los estudios universitarios en nuestro continente. En este sentido, un sistema de créditos específico, el sistema de ECTS (siglas en inglés de Sistema Europeo de Transferencia de Créditos), fue diseñado para describir un programa educativo. El ECTS es un sistema centrado en el estudiante que se basa en la carga de trabajo requerida para alcanzar los objetivos del programa. Así, cada centro proporciona una descripción de los cursos junto con su valor en créditos para facilitar la selección de los mismos a los estudiantes de otras universidades. Los estudiantes obtienen de esta forma el reconocimiento de los estudios cursados en otras instituciones con las que su universidad de destino mantiene acuerdos de colaboración transfiriendo así sus créditos académicos de una institución a otra.

A los cambios en el sistema universitario europeo hay que añadir la actual implicación de los jóvenes con las nuevas tecnologías. En particular, según un estudio realizado por Telefónica, los usuarios de Internet en España se concentran en el rango de edades entre 16 y 34, son estudiantes y han adoptado Internet como herramienta indispensable de trabajo y ocio<sup>1</sup>. De los usuarios que semanalmente se conectan a Internet, el 77% son estudiantes<sup>2</sup>. El aprendizaje a través de la red es en estos momentos una posibilidad para los estudiantes y comienza a utilizarse en la formación universitaria.

La integración de estas tecnologías y los programas de colaboración entre universidades en el proceso de enseñanza universitario es pues una realidad que requiere de un proceso de análisis y experimentación. Así, este artículo revisa los principales métodos de enseñanza-aprendizaje susceptibles de aplicación en los procesos de colaboración interuniversitarios. A partir de ahí, se describe el proyecto interdisciplinar y bilingüe desarrollado entre dos centros presentando los principales resultados relativos a la participación y rendimiento de los estudiantes implicados.

## Métodos docentes

El desarrollo de un novedoso, por ser interdisciplinar, bilingüe y virtual, proyecto docente requiere de una obligada revisión del proceso de contenidos relativos al pro-

<sup>(1)</sup> <http://www.Internetnews.com/bus-news/article.php/713971>

<sup>(2)</sup> <http://www.lukor.com/ordenadores/06040602.htm>

ceso enseñanza-aprendizaje así como de los principales métodos de enseñanza susceptibles de ser empleados.

En primer lugar, la *enseñanza* se refiere al conjunto de decisiones, actividades y medios que se organizan sistemáticamente para facilitar el aprendizaje del estudiante (De la Cruz, 1994). Y por su parte, el *aprendizaje* hace referencia a adquisiciones o cambios relativamente estables en la conducta y los conocimientos del aprendiz como efecto de un programa de instrucción. Se trata, como se ha señalado de manera concluyente, de dos estados interrelacionados, como señalan Brown y Atkins (1988), de forma que la función básica y esencial de la enseñanza es el aprendizaje (Novack, 1982). En este proceso, resulta necesario utilizar métodos y recursos didácticos que permitan desarrollar satisfactoriamente el proceso de enseñanza-aprendizaje (Ortega, 1992).

La existencia de múltiples métodos de enseñanza ha llevado a su clasificación por distintos autores. Pujol y Fons (1981) definen las formas didácticas que consideran básicas en la enseñanza universitaria: la lección magistral, la enseñanza en pequeños grupos y el sistema tutorial. Otros autores, Brown y Atkins (1988) mantienen las dos primeras categorías, la lección magistral y la enseñanza en grupos pequeños, presentando dos nuevas categorías: la enseñanza en laboratorios y la investigación independiente con supervisor. De manera similar, Colom et al. (1988) clasifican los métodos de enseñanza en cuatro principales grupos: métodos de enseñanza colectiva, en grupo, individualizada, y de enseñanza experimental. Finalmente y desde una perspectiva más amplia, Fernández, Sarramora y Tarín (1984) determinan tres modalidades de docencia: docencia magistral, socializada e individual. De todos esos métodos de enseñanza, las clases prácticas requieren de aquéllos que permitan al estudiante adquirir habilidades necesarias para su ejercicio profesional, mediante el desarrollo de capacidades analíticas, de razonamiento, de resolución de problemas, y de comunicación, entre otras. Así, estas clases se desarrollarán mediante la interacción y participación activa de los estudiantes y no a través de métodos meramente expositivos. En este contexto, consideramos oportuno emplear los métodos de *investigación independiente con supervisor*, a través de la discusión en pequeños grupos (Pujol y Fons, 1981). Estos métodos permiten al alumno trabajar en equipo, pudiendo de esta manera presentar y defender sus aportaciones, resolver problemas, aplicar conocimientos, evaluar las ideas de los demás, y recibir confirmación de su proceso de asimilación, entre otros aspectos. En definitiva permiten, según Pujol y Fons (1981), lograr la individualización de la enseñanza, conseguir la participación activa de todos los alumnos y aprender a trabajar en equipo.

Según la mayor o menor formalidad y estructura, los anteriores autores clasifican las técnicas a desarrollar en estos grupos en tres tipos: *técnicas más estructuradas*,

como el simposio, la mesa redonda, el panel, el examen por una comisión y el debate público; *técnicas menos estructuradas o más informales*, como reuniones preliminares en pequeños grupos y diálogos simultáneos; y *otras técnicas*, como el método del caso, las actividades de representación o *role-playing* y los juegos de empresa.

De las anteriores, adquieren un papel destacado aquellas menos estructuradas o más informales, principalmente si se parte del supuesto de que para llegar al nivel de «dominio» de conocimientos, destrezas y actitudes, el camino más corto e interesante es el aprendizaje activo (aprender haciendo) y cooperativo (aprender con los pares). Según Slavin (1990), el aprendizaje cooperativo se define como estrategias de instrucción en las que los estudiantes trabajan divididos en pequeños grupos en actividades de aprendizaje y son evaluados según la productividad del grupo.

En este contexto, la utilización de medios y recursos didácticos en las distintas metodologías docentes resulta sin duda de gran ayuda. Los *medios y recursos didácticos*, según señala Cebrián (1994), constituyen una parte de todo programa docente al determinar a los otros elementos (objetivos, contenidos, metodologías) y ser orientados, diseñados y producidos en función de éstos. En este sentido, considera que los medios y recursos didácticos:

- Posibilitan al alumnado y profesorado *informaciones y datos* que escapan a las variables temporales o espaciales.
- Permiten *simular experiencias* difíciles de realizar.
- Ofrecen la posibilidad de *desarrollar habilidades y capacidades cognitivas y afectivas*.
- Provocan *modelos de comunicación específicos* entre las personas que viven los procesos de enseñanza-aprendizaje.
- Facilitan la *transmisión, representación y reconstrucción* de los contenidos académicos y añaden nuevos elementos de estudio.
- Pueden generar *procesos de autoevaluación* del alumnado y del profesorado, bajo el análisis, estudio y reflexión de datos almacenados y recogidos por éstos.

Los medios y recursos didácticos, si bien necesarios, deben adquirir un papel no protagonista en nuestras aulas, pues las relaciones humanas deben ser el eje fundamental de la educación y el papel del profesor no puede ser desbancado por estos, aunque a veces ganen en eficacia.

Los recursos que pueden utilizarse para presentar los contenidos de la materia pueden ser de dos tipos: impresos (manuales de la materia, libros especializados,

libros de ejercicios y casos, artículos de revistas académicas y profesionales, ponencias, artículos de prensa, etc.) y audiovisuales. Dentro de estos últimos, de mayor impacto en el aula por sus notorias condiciones de comunicación, se encuentra el uso de programas informáticos y de la conexión a Internet. En esta línea, los importantes avances en aplicaciones multimedia y el desarrollo de las telecomunicaciones, permiten desarrollar nuevas alternativas didácticas orientadas a facilitar el proceso de enseñanza-aprendizaje (Sandoval y García, 2000; García, 2002; Mir et al., 2003).

Todas estas técnicas didácticas permiten, tal y como señalaban Pujol y Fons (1981), entre otras ventajas, promover la interacción en el proceso de enseñanza-aprendizaje, no limitándose a cubrir una labor exclusivamente de transmisión sino construir el conocimiento en nuestras aulas. En síntesis, el empleo conjunto de todo el material disponible es especialmente recomendable pues la utilización de más de un canal facilita la percepción en el proceso de comunicación y refuerza el proceso de enseñanza-aprendizaje (Fernández et al., 1984).

## Proyecto docente interdisciplinar

En este entorno y consecuencia de la firma de un convenio de colaboración firmado entre la London School of Economics (LSE) del Reino Unido y la Universidad de Valencia (UV) de España, se diseñó un proyecto interdisciplinar, bilingüe y virtual de enseñanza-aprendizaje (Byrne et al., 2007). El proyecto, realizado a través de una plataforma virtual (webCT), se desarrolló en español para los temas propuestos por la LSE y en inglés para los programados por la UV. Una colaboración basada no sólo en aspectos lingüísticos, sino también en cuestiones de economía y gestión.

La LSE, creada en 1895, no sólo ofrece cursos de *Economía y Ciencias Políticas*, sino también un amplio abanico de asignaturas de «Ciencias Sociales». Además, su renombrado Centro de Idiomas ofrece diversas asignaturas optativas a los estudiantes de grado, entre las que se encuentra la asignatura de «Lengua Española y Sociedad». Asignatura cursada por estudiantes no castellano-hablantes. La UV, una de las universidades más antiguas de Europa, ofrece en la actualidad una amplia gama de titulaciones, entre las que se encuentra la de *Administración y Dirección de Empresas*, donde se imparte la asignatura troncal «Dirección Comercial II». La docencia de esta asignatura se oferta en castellano, valenciano e inglés pudiendo el alumno elegir en

función de sus intereses. El grupo en inglés es sobre el que se desarrolló el proyecto interdisciplinar de colaboración entre ambas instituciones, no siendo este idioma la lengua materna de la casi totalidad de sus estudiantes.

El objetivo del proyecto era introducir nuevos recursos didácticos en la docencia de dos materias diferentes de dos universidades de distintos países. De este modo, se producía una colaboración interdisciplinaria entre un centro de idiomas y un departamento académico, por medio de la cual los estudiantes intercambiaran conocimientos y experiencias en dos lenguas diferentes a través de un entorno virtual de aprendizaje (EVA).

Los objetivos del proyecto fueron los siguientes:

- Promover, a través el uso de nuevas tecnologías, la cooperación y el intercambio de información entre los estudiantes de «Español y Sociedad» en LSE y los estudiantes de «Administración de Empresas» en inglés en la UV.
- Implicar a los estudiantes en actividades que desarrollen sus habilidades lingüísticas en el idioma en el que cursan la materia así como sus actitudes, valores y creencias de otras realidades o culturas.
- Crear un ambiente único y provechoso para discutir temas relacionados con las ciencias sociales (esto es, política, historia, economía, gestión de empresas o comercialización).
- Motivar el aprendizaje autónomo en un determinado tema.
- Promover el pensamiento crítico y las habilidades analíticas al comparar aspectos culturales, sociales y de gestión de negocios por medio de tareas específicas.
- Desarrollar habilidades transferibles tales como técnicas de negociación y gestión del tiempo.

Con todo, este trabajo presenta tal experiencia de colaboración y explora la conveniencia del entorno virtual para estudiantes de distintas disciplinas, instituciones y países.

El proyecto comenzó en febrero de 2006 y se desarrolló a lo largo de seis semanas. Durante este periodo, los estudiantes participaron en diversas actividades orales y escritas y en discusiones guiadas por los profesores que coordinaban el proyecto, todos ellos profesores de las asignaturas en cuestión. Todos los intercambios debían ser realizados a través de la plataforma de enseñanza virtual webCT a la que todos los estudiantes tenían acceso. Los intercambios de información entre los estudiantes fueron supervisados por los profesores responsables solicitándose a los estudiantes

determinadas tareas relacionadas con las actividades (trabajos puntuales, informes, presentaciones orales o proyecto de investigación personal). Para las primeras dos tareas, se agrupó a los estudiantes en equipos de cuatro personas para facilitar el intercambio de la información.

Los estudiantes de la LSE tuvieron asimismo ocasión de visitar la universidad colaboradora durante un breve viaje de estudios organizado por ambas instituciones. Durante su visita los estudiantes pudieron conocer a los estudiantes y profesores participantes en el proyecto, asistir a diversas clases y conferencias, y participar en actividades extracurriculares tales como reuniones con representantes políticos periodistas, y visitas a instituciones políticas y culturales.

## Participación y resultados

En el proyecto participaron 45 estudiantes desde la LSE y 32 desde la UV. La participación de los estudiantes se contabilizó a través del número de accesos a la plataforma virtual, el número de mensajes leídos, el número de mensajes enviados y la fecha de la primera y de la última conexión a la plataforma. Los resultados anteriores se analizaron junto con las calificaciones obtenidas en el examen de las respectivas asignaturas una vez finalizó el proyecto.

Los datos de participación de los estudiantes mostraron el gran interés de los mismos ya que el nivel de uso fue más elevado de lo esperado. El número de conexiones totales fue de 587 en solamente cuatro semanas. Ello nos lleva a afirmar que la interacción entre los estudiantes de la LSE y la UV fue una realidad. Además, el análisis cualitativo de los diversos informes que los estudiantes presentaron a lo largo del curso evidencia que fueron capaces de utilizar y estructurar la información proporcionada por sus colaboradores para las preguntas planteadas.

A nivel cuantitativo, para los estudiantes de la UV, la Tabla I muestra información de accesos (primer y último acceso, impactos, número de mensajes leídos y número de enviados) así como la nota final (que puede tomar valores entre 0 y 10). Si bien aparentemente parece que algunos estudiantes no han participado en el proyecto, dado que sus impactos son nulos, existe la posibilidad de que trabajasen en grupo y participasen utilizando únicamente la contraseña de uno de los miembros del equipo.

TABLA I. Accesos y resultados para los estudiantes de la UV

Estudiante	Primer acceso		Último acceso		Impactos	Leídos	Enviados	Nota
1	February 15, 2006	4:51pm	April 6, 2006	2:55pm	134	55	23	7.6
2	February 16, 2006	1:22pm	March 9, 2006	1:07pm	79	38	8	1.3
3	February 16, 2006	1:22pm	March 23, 2006	6:22pm	116	95	10	0.6
4	February 16, 2006	1:22pm	April 27, 2006	12:48pm	68	42	7	2.2
5	February 16, 2006	1:22pm	March 22, 2006	11:44am	27	23	0	NA
6	February 16, 2006	1:22pm	March 30, 2006	9:24am	104	66	16	8.1
7	February 16, 2006	1:22pm	April 27, 2006	12:51pm	41	16	7	5.5
8	February 16, 2006	1:22pm	April 27, 2006	12:50pm	149	113	19	2.0
9	February 16, 2006	1:22pm	April 27, 2006	12:47pm	158	113	8	2.6
10	February 16, 2006	1:22pm	April 6, 2006	12:54pm	51	18	7	2.2
11	February 16, 2006	1:22pm	May 5, 2006	9:17am	113	45	15	6.1
12	February 16, 2006	1:23pm	March 2, 2006	8:27pm	79	40	8	6.2
13	February 16, 2006	1:32pm	June 1, 2006	5:59pm	199	76	24	2.3
14	February 16, 2006	1:37pm	April 27, 2006	12:55pm	42	28	1	5.8
15	February 19, 2006	10:18pm	May 17, 2006	8:54am	128	72	18	5.2
16	February 20, 2006	5:20pm	April 27, 2006	12:48pm	118	48	23	3.0
17	February 20, 2006	6:05pm	April 6, 2006	12:52pm	69	42	9	2.1
18	February 21, 2006	10:39am	February 21, 2006	10:39am	2	0	0	5.7
19	February 23, 2006	12:45pm	April 27, 2006	12:48pm	77	46	7	6.5
20	February 23, 2006	12:48pm	April 27, 2006	12:46pm	176	102	29	1.1
21	March 2, 2006	12:59pm	April 27, 2006	12:47pm	51	35	3	5.2
22	March 22, 2006	6:10pm	March 22, 2006	6:13pm	5	2	0	1.6
23	March 9, 2006	12:46pm	March 11, 2006	1:02pm	42	32	4	2.4
24	March 9, 2006	12:51pm	April 27, 2006	1:01pm	80	63	1	1.3
25	---	---	---	---	0	0	0	0.9
26	---	---	---	---	0	0	0	1.3
27	---	---	---	---	0	0	0	2.3
28	---	---	---	---	0	0	0	1.4
29	---	---	---	---	0	0	0	2.2
30	---	---	---	---	0	0	0	NA
31	---	---	---	---	0	0	0	0.8
32	---	---	---	---	0	0	0	NA

Para determinar la correlación entre participación e implicación de los estudiantes en el proyecto de *e-learning* y su rendimiento académico en el curso, utilizamos el coeficiente de concordancia por rangos de Spearman. Este coeficiente mide el grado y el sentido de la relación entre el *ranking* de los estudiantes en cuanto a su grado de implicación y su posición en el *ranking* de notas.

**TABLA II. Coeficiente de concordancia por rangos de Spearman entre los indicadores de participación y las notas finales de la asignatura para los estudiantes de la UV**

Indicadores de participación	Coeficiente de Spearman
Impactos	0.6373*
Mensajes leídos	0.5766*
Mensajes enviados	0.7289*
Primera conexión	0.6910*
Duración del periodo de conexión	0.7410*

(\*) Estadísticamente significativo al 1%.

Como se puede observar, todos los coeficientes de concordancia por rangos obtenidos muestran valores positivos y significativos al 1%, es decir, existe una relación positiva fuerte entre la implicación y la participación de estos estudiantes en las actividades del proyecto de aprendizaje virtual y sus resultados finales en la asignatura. En concreto, cuanto mayor es el número de mensajes enviados y cuanto más extensa es la duración del período de conexión, mejores son los resultados obtenidos por el estudiante. Estos dos indicadores, es decir, su participación activa y la continuidad de este esfuerzo a lo largo del tiempo, son los que implican mayor esfuerzo para el estudiante. Esta evidencia parece apoyar la influencia positiva de esta actividad interdisciplinaria en el rendimiento de los estudiantes de la UV en la asignatura de «Dirección Comercial II».

Por otra parte, para los estudiantes de la LSE, la Tabla III muestra información de accesos (primer y último acceso, impactos, número de mensajes leídos y enviados) así como la nota final (que puede tomar valores entre 0% y 100%).

**TABLA III. Accesos y resultados para los estudiantes de la LSE**

Estudiante	Primer acceso		Último acceso		Impactos	Leídos	Enviados	Nota
1	February 16, 2006	1:21pm	April 9, 2006	1:01am	172	118	10	71%
2	February 16, 2006	1:56pm	May 11, 2006	1:53pm	122	80	16	69%
3	February 16, 2006	12:49pm	February 27, 2006	2:14pm	29	19	3	75%
4	February 16, 2006	12:18pm	May 12, 2006	10:56am	61	16	15	58%
5	February 16, 2006	1:22pm	February 27, 2006	7:03pm	7	0	0	0%
6	---	---	---	---	0	0	0	---
7	February 26, 2006	1:15pm	February 26, 2006	1:16pm	4	2	0	---
8	February 16, 2006	1:21pm	April 1, 2006	9:36pm	61	35	5	0%
9	February 16, 2006	12:17pm	April 19, 2006	11:51pm	128	57	15	67%
10	April 9, 2006	10:13am	April 9, 2006	10:13am	2	0	0	---
11	February 23, 2006	1:11pm	April 16, 2006	3:50pm	149	101	8	70%
12	February 16, 2006	1:21pm	May 12, 2006	11:51am	243	152	17	66%
13	March 21, 2006	10:11am	March 21, 2006	10:14am	1	0	0	---
14	---	---	---	0	0	0	71%	---
15	February 24, 2006	5:52pm	July 11, 2006	6:35am	156	77	6	66%

Estudiante	Primer acceso		Último acceso		Impactos	Leídos	Enviados	Nota
16	February 16, 2006	12:17pm	February 16, 2006	12:50pm	12	2	6	53%
17	February 16, 2006	12:18pm	April 5, 2006	4:33pm	106	62	17	63%
18	April 3, 2006	2:48pm	April 3, 2006	2:48pm	1	0	0	
19	February 16, 2006	1:22pm	May 12, 2006	1:28pm	107	52	16	73%
20	April 26, 2006	4:49pm	April 26, 2006	4:51pm	2	0	0	
21	October 8, 2006	2:22am	October 8, 2006	2:22am	1	0	0	
22	February 16, 2006	12:17pm	February 23, 2006	12:19pm	17	8	1	64%
23	February 16, 2006	12:18pm	April 30, 2006	11:54am	97	83	0	66%
24	February 16, 2006	1:22pm	May 4, 2006	6:23pm	151	69	13	69%
25	---	---	---	---	0	0	0	
26	February 16, 2006	1:28pm	March 27, 2006	11:21pm	80	44	17	
27	February 15, 2006	3:02pm	February 15, 2006	3:02pm	3	0	0	
28	February 16, 2006	1:26pm	May 17, 2006	9:54am	231	133	33	70%
29	February 16, 2006	1:22pm	February 23, 2006	1:56pm	53	34	14	62%
30	---	---	---	---	0	0	0	67%
31	July 23, 2006	6:30pm	July 23, 2006	6:30pm	1	0	0	
32	February 23, 2006	12:11pm	April 24, 2006	11:39am	60	35	13	68%
33	February 16, 2006	1:25pm	May 20, 2006	5:35pm	135	61	19	71%
34	July 2, 2006	1:09pm	July 2, 2006	1:09pm	1	0	0	
35	February 16, 2006	1:23pm	May 6, 2006	6:31pm	79	54	6	64%
36	---	---	---	---	0	0	0	
37	February 16, 2006	1:22pm	May 5, 2006	9:17am	113	45	15	64%
38	February 16, 2006	12:19pm	March 2, 2006	12:55pm	47	23	5	59%
39	May 20, 2006	1:34pm	May 20, 2006	1:34pm	3	0	0	
40	February 16, 2006	1:28pm	September 6, 2006	12:28pm	131	79	18	
41	---	---	---	---	0	0	0	
42	February 16, 2006	12:17pm	March 21, 2006	5:05pm	145	88	6	65%
43	February 16, 2006	12:19pm	February 16, 2006	12:51pm	10	2	7	61%
44	February 21, 2006	10:39am	February 21, 2006	10:39am	2	0	0	76%
45	February 16, 2006	12:18pm	April 30, 2006	11:19am	114	55	9	68%

Al igual que para el caso español, con el fin de evaluar el grado de relación entre la participación de los estudiantes de la LSE en el proyecto de aprendizaje virtual y sus resultados en la asignatura, se estima el coeficiente de concordancia por rangos de Spearman. Los resultados obtenidos se muestran en la Tabla IV.

TABLA IV. Coeficiente de concordancia por rangos de Spearman entre los indicadores de participación y las notas finales de la asignatura para los estudiantes de la LSE

Indicadores de participación	Coeficiente de Spearman
Impactos	0.9346*
Mensajes leídos	0.7604*
Mensajes enviados	0.8859*
Primera conexión	0.8628*
Duración del periodo de conexión	0.9023*

(\*) Estadísticamente significativo al 1%.

De forma similar al caso de la UV, todos los coeficientes de concordancia obtenidos para los estudiantes de la LSE arrojan valores positivos y significativos al 1%, corroborando también la existencia de una relación positiva fuerte entre la participación y el rendimiento académico del estudiante en la asignatura «Lengua Española y Sociedad».

## Conclusión

Los resultados obtenidos para ambas universidades parecen apoyar la influencia positiva de esta actividad interdisciplinaria en el rendimiento de los estudiantes y, con ello, la utilización de métodos de enseñanza que faciliten el aprendizaje activo y cooperativo por medio de recursos didácticos audiovisuales, en línea con los planteamientos de de Pujol y Fons (1981) y Cebrián (1994).

Sin embargo, los coeficientes de concordancia por rangos para la LSE son superiores a los obtenidos para la UV. Esto puede llevarnos a pensar que la experiencia ha contribuido a mejorar en mayor medida los resultados de los estudiantes de «Lengua Española y Sociedad» que para los estudiantes de *Dirección Comercial II*. Si bien este resultado puede obedecer a distintos factores, la conveniencia de este tipo de colaboración en la enseñanza del español como segunda lengua para los estudiantes no castellano-hablantes parece fuera de toda duda.

En conclusión, este proyecto de colaboración permitió a estudiantes de distintas instituciones y disciplinas no sólo conocer a otros estudiantes de diferentes titulaciones sino también debatir, obtener y procesar información sobre otras realidades, así como mejorar sus habilidades en otro idioma a través de una plataforma virtual.

Los resultados obtenidos para los estudiantes de ambas instituciones -Universidad de Valencia y London School of Economics- en esta actividad de aprendizaje virtual son positivos tanto en términos cuantitativos como cualitativos. Sin embargo, la evidencia no es concluyente, ya que se trata de un proyecto experimental y la actividad se debe replicar en otros cursos y asignaturas.

Entendemos además que nuestro trabajo no está exento de limitaciones. En cuanto a los datos cuantitativos, los resultados obtenidos podrían estar sesgados debido a que los estudiantes trabajan en grupos utilizando en ocasiones solamente la contraseña de uno de los miembros. Por otra parte, la causalidad entre

los indicadores de participación y de rendimiento académico final del estudiante no ha sido explorada. En este sentido, cabe preguntarse si los estudiantes tienen mejores resultados gracias a su participación en esta actividad de aprendizaje virtual o son los «mejores» estudiantes los que tienden a participar en esta actividad en mayor medida que los demás. Clarificar esta última cuestión, requiere de investigación adicional considerando el rendimiento de los estudiantes en otras asignaturas.

En términos generales, los resultados parecen confirmar la contribución positiva de esta actividad interdisciplinaria, bilingüe y virtual al proceso enseñanza-aprendizaje.

## Referencias bibliográficas

- BROWN, G. & ATKINS, M. (1988). *Effective teaching in higher education*. London: Methue & Ltd.
- BYRNE, N., COCA, M. Y CUADRADO, M. (2007). Interdisciplinary university e-learning through weblab systems. A two-country project. INTED Conference, Valencia, Spain.
- CEBRIÁN, M. (1994). Medios didácticos. En SERVEI DE FORMACIÓ PERMANENT (ed.), *Medios y recursos didácticos en el ámbito universitario* (pp. 1-29). Universitat de València.
- COLOM, A., SUREDA, J. Y MARTÍN, J. (1988). *Tecnología y medios educativos*. Madrid: Cincel, Serie Educación y futuro, 3.
- DE LA CRUZ, M. A. (1994). *Evaluación del conocimiento y su adquisición*. En R. FERNÁNDEZ (ed.), *Introducción a las técnicas de evaluación psicológica*. Madrid: Pirámide.
- FERNÁNDEZ, A., SARRAMORA, J. Y TARÍN, L. (1984). *Tecnología didáctica. Teoría y práctica de la programación escolar*. Barcelona: CEAC.
- GARCÍA, L. (2002). *La educación a distancia: De la teoría a la práctica*. Madrid: Ariel.
- MIR, J. I., REPARAZ, C. Y SOBRINO, A. (2003). *La formación en Internet. Modelo de un curso on line*. Madrid: Ariel.
- NOVACK, J. D. (1982). *Teoría y práctica de la educación*. Madrid: Alianza.
- ORTEGA, E. (1992). La enseñanza del Marketing y el reto del futuro. *ESIC Market* 75, 9-28.
- PUJOL, J. Y FONS, J. L. (1981). *Los métodos en la enseñanza universitaria*. Pamplona: Ediciones Universidad de Navarra.

- SANDOVAL, A. Y GARCÍA, A. (2000). Las nuevas tecnologías y la docencia universitaria: desarrollo de un sistema de soporte a la actividad docente e investigadora del departamento de Organización de Empresas y Marketing de la Universidad de Vigo. La inteligencia empresarial: la gestión del conocimiento en la empresa. Actas del XIV Congreso Nacional y X Congreso Hispano-francés, Universidad de Jaén.
- SLAVIN, R. E. (1990). *Cooperative learning*. New Jersey: Prentice-Hall.

## Fuentes electrónicas

- EUROPEAN COMMISSION (2007). Consultado el 1 de marzo de 2007, de: [http://ec.europa.eu/education/programmes/socrates/ects/index\\_en.html](http://ec.europa.eu/education/programmes/socrates/ects/index_en.html) .
- EUROPA PRESS (2000) El 50% de los españoles no ha utilizado nunca Internet. *Lokuor*: Consultado de: <http://www.lukor.com/ordenadores/06040602.htm>
- INTERNETWORLDSTAT (2007). Consultado el 1 de marzo de 2007, de: <http://www.Internetworldstats.com>.
- LONDON SCHOOL OF ECONOMICS (2007). Consultado el 1 de marzo de 2007, de: <http://www.lse.ac.uk> .
- Spain's Broadband Future Loiters in Cybercafes (2001, 15 de marzo). *InternetNews*. Consultado de: <http://www.Internetnews.com/bus-news/article.php/713971>
- UNIVERSITAT DE VALÈNCIA (2007). Consultado el 1 de marzo de 2007, de: <http://www.lse.ac.uk>.

**Dirección de contacto:** M<sup>a</sup> Eugenia Ruiz. Universidad de Valencia. Facultad de Economía. Avda de los Naranjos, s/n. 46022, Valencia, España. E-mail: M.Eugenia.Ruiz@uv.es



## **Índice bibliográfico**



## **Índice bibliográfico de la *Revista de Educación (Madrid)*. Año 2008 (números 345-346-347 y extraordinario)**

## **Bibliographical Index of *Revista de Educación (Madrid)*. Year 2008 (Issues 345-346-347 and Extraordinary)**

La *Revista de Educación (Madrid)* pone a disposición de sus lectores un índice bibliográfico que facilita la consulta de todos los artículos y reseñas publicados durante el último año. Este índice recoge los tres números ordinarios (345, 346 y 347) y el extraordinario del año 2008. En primer lugar, se relacionan los artículos que componen la sección monográfica de la Revista, clasificados y ordenados por números. A continuación, se presentan los artículos publicados en la sección ordinaria dividida en tres subsecciones: *Investigaciones y estudios*, *Ensayos e informes* y *Experiencias educativas (innovación)*, todos ellos ordenados alfabéticamente. En tercer lugar, se muestran los artículos del número extraordinario dedicado a los tiempos de cambio universitario en Europa. Por último, se presenta una relación de los libros reseñados durante el 2008, así como una lista alfabética de los autores que han publicado en estos cuatro números de la Revista.

Este índice puede consultarse en la página web: [www.revistaeducacion.mec.es](http://www.revistaeducacion.mec.es).

*Revista de Educación (Madrid)* puts at the disposal of its readers a bibliographical index that facilitates the consultation of all the papers and reviews published during the last year. This index contains the three regular issues and extraordinary one along 2008. Firstly, there is a list of the papers of the monograph section of the Journal, classified in the same order as they are in the issue. After that, the papers published in the regular section divided in three subsections: *Research and Studies*, *Reports and Essays and Education Experiences (Innovation)*, in alphabetical order. Thirdly, the papers from the extraordinary issue dedicated to times of university change in Europa are listed. Finally, there is a list of the reviewed books in 2008 and a list of all the authors who have published in 2008, too.

This index is available in the web of the Journal: [www.revistaeducacion.mec.es](http://www.revistaeducacion.mec.es)



# Índice bibliográfico de la Revista de Educación (Madrid). Año 2008 (números 345-346-347 y extraordinario)

## Bibliographical Index of Revista de Educación (Madrid). Year 2008 (Issues 345-346-347 and Extraordinary)

### Sección monográfica

**Número 345.** De inmigrantes a minorías: temas y problemas de la multiculturalidad  
(Coordinador: Mariano Fernández Enguita)

FERNÁNDEZ ENGUITA, M. Y TERRÉN, E.

Presentación

En: *Revista de Educación*, Ministerio de Educación, Política Social y Deporte, 345, 2008, pp. 15-21. [[http://www.revistaeducacion.mec.es/re345\\_01.html](http://www.revistaeducacion.mec.es/re345_01.html)]. [ISSN 0034-8082]

GARCÍA CASTAÑO, F.J., RUBIO GÓMEZ, M. Y BOUACHRA, O.

Población inmigrante y escuela en España: un balance de investigación

En: *Revista de Educación*, Ministerio de Educación, Política Social y Deporte, 345, 2008, pp. 23-60. [[http://www.revistaeducacion.mec.es/re345\\_02.html](http://www.revistaeducacion.mec.es/re345_02.html)]. [ISSN 0034-8082]

ALEGRE CANOSA, M.A.

Educación e inmigración ¿Un binomio problemático?

En: *Revista de Educación*, Ministerio de Educación, Política Social y Deporte, 345, 2008, pp. 61-82. [[http://www.revistaeducacion.mec.es/re345\\_03.html](http://www.revistaeducacion.mec.es/re345_03.html)]. [ISSN 0034-8082]

TORRES SANTOMÉ, J.

Diversidad cultural y contenidos escolares

En: *Revista de Educación*, Ministerio de Educación, Política Social y Deporte, 345, 2008, pp. 83-110. [[http://www.revistaeducacion.mec.es/re345\\_04.html](http://www.revistaeducacion.mec.es/re345_04.html)]. [ISSN 0034-8082]

FRANZÉ MUDANÓ, A.

Diversidad cultural en la escuela. Algunas contribuciones antropológicas

En: *Revista de Educación*, Ministerio de Educación, Política Social y Deporte, 345, 2008, pp. 111-132. [[http://www.revistaeducacion.mec.es/re345\\_05.html](http://www.revistaeducacion.mec.es/re345_05.html)]. [ISSN 0034-8082]

GARRETA BOCHACA, J.

Escuela, familia de origen inmigrante y participación

En: *Revista de Educación*, Ministerio de Educación, Política Social y Deporte, 345, 2008, pp. 133-155. [[http://www.revistaeducacion.mec.es/re345\\_06.html](http://www.revistaeducacion.mec.es/re345_06.html)]. [ISSN 0034-8082]

FERNÁNDEZ ENGUITA, M., GAETE, J.M. Y TERRÉN, E.

¿Fronteras en las aulas? Contacto transcultural y endogamia en las interacciones del alumnado

En: *Revista de Educación*, Ministerio de Educación, Política Social y Deporte, 345, 2008, pp. 157-181. [[http://www.revistaeducacion.mec.es/re345\\_07.html](http://www.revistaeducacion.mec.es/re345_07.html)]. [ISSN 0034-8082]

CARRASCO CARPIO, C. Y RIESCO SANZ, A.

Escuela, consumo y el mercado de trabajo: la producción de la «juventud» entre los jóvenes de origen inmigrante

En: *Revista de Educación*, Ministerio de Educación, Política Social y Deporte, 345, 2008, pp. 183-203. [[http://www.revistaeducacion.mec.es/re345\\_08.html](http://www.revistaeducacion.mec.es/re345_08.html)]. [ISSN 0034-8082]

**Número 346.** El análisis de la interacción alumno-profesor: líneas de investigación  
(Coordinador: César Coll)

COLL, C. Y SÁNCHEZ, E.

Presentación

En: *Revista de Educación*, Ministerio de Educación, Política Social y Deporte, 346, 2008, pp. 15-32. [[http://www.revistaeducacion.mec.es/re346\\_01.html](http://www.revistaeducacion.mec.es/re346_01.html)]. [ISSN 0034-8082]

COLL, C., ONRUBIA, J. Y MAURI, T.

Ayudar a aprender en contextos educativos: el ejercicio de la influencia educativa y el análisis de la enseñanza

En: *Revista de Educación*, Ministerio de Educación, Política Social y Deporte, 346, 2008, pp. 33-70. [[http://www.revistaeducacion.mec.es/re346\\_02.html](http://www.revistaeducacion.mec.es/re346_02.html)]. [ISSN 0034-8082]

CUBERO PÉREZ, R., CUBERO PÉREZ, M., SANTAMARÍA SANTIGOSA, A., DE LA MATA BENÍTEZ, M. L., IGNACIO CARMONA, M. J. Y PRADOS GALLARDO, M. M.

La educación a través de su discurso. Prácticas educativas y construcción discursiva del conocimiento en el aula

En: *Revista de Educación*, Ministerio de Educación, Política Social y Deporte, 346, 2008, pp. 71-104. [[http://www.revistaeducacion.mec.es/re346\\_03.html](http://www.revistaeducacion.mec.es/re346_03.html)]. [ISSN 0034-8082]

SÁNCHEZ, E., GARCÍA, J. R., ROSALES, J., DE SIXTE, R Y CASTELLANO, N.

Elementos para analizar la interacción entre estudiantes y profesores: ¿qué ocurre cuando se consideran diferentes dimensiones y diferentes unidades de análisis?

En: *Revista de Educación*, Ministerio de Educación, Política Social y Deporte, 346, 2008, pp. 105-136. [[http://www.revistaeducacion.mec.es/re346\\_04.html](http://www.revistaeducacion.mec.es/re346_04.html)]. [ISSN 0034-8082]

**Número 347.** Los derechos de la primera infancia (0-6 años): atención socioeducativa (Coordinadora: María Paz Lebrero)

Lebrero Baena, M<sup>a</sup>. P.

Presentación.

En: *Revista de Educación*, Ministerio de Educación, Política Social y Deporte, 347, 2008, pp. 15-31. [[http://www.revistaeducacion.mec.es/re347\\_01.html](http://www.revistaeducacion.mec.es/re347_01.html)]. [ISSN 0034-8082]

BLANCO GUIJARRO, R.

Construyendo las bases de la inclusión y la calidad de la educación en la primera infancia.

En: *Revista de Educación*, Ministerio de Educación, Política Social y Deporte, 347, 2008, pp. 33-54. [[http://www.revistaeducacion.mec.es/re347\\_02.html](http://www.revistaeducacion.mec.es/re347_02.html)]. [ISSN 0034-8082]

GUIJO BLANCO, V.

Derechos de la infancia (0-6 años) y educación inclusiva.

En: *Revista de Educación*, Ministerio de Educación, Política Social y Deporte, 347, 2008, pp. 55-74. [[http://www.revistaeducacion.mec.es/re347\\_03.html](http://www.revistaeducacion.mec.es/re347_03.html)]. [ISSN 0034-8082]

SÁNCHEZ SÁNCHEZ, A.

Enfoque, estabilidad y continuidad de la escolarización en Educación Infantil (0-6 años). Evolución histórica en la Comunidad de Madrid

En: *Revista de Educación*, Ministerio de Educación, Política Social y Deporte, 347, 2008, pp. 75-99. [[http://www.revistaeducacion.mec.es/re347\\_04.html](http://www.revistaeducacion.mec.es/re347_04.html)]. [ISSN 0034-8082]

PINEDA HERRERO, P., MORENO ANDRÉS, M<sup>a</sup>. V., ÚCAR MARTÍNEZ, X. Y BELVIS PONS, E.

Derecho a la calidad: evaluación de la formación permanente en el sector de la Educación Infantil en España

En: *Revista de Educación*, Ministerio de Educación, Política Social y Deporte, 347, 2008, pp. 101-126. [[http://www.revistaeducacion.mec.es/re347\\_05.html](http://www.revistaeducacion.mec.es/re347_05.html)]. [ISSN 0034-8082]

HORNO GOICOECHEA, P.

Salvaguardar los derechos desde la escuela: educación afectivo-sexual para la prevención primaria del maltrato infantil

En: *Revista de Educación*, Ministerio de Educación, Política Social y Deporte, 347, 2008, pp. 127-140. [[http://www.revistaeducacion.mec.es/re347\\_06.html](http://www.revistaeducacion.mec.es/re347_06.html)]. [ISSN 0034-8082]

VALCARCE GARCÍA, A.

Políticas Públicas para la Atención Temprana de la Infancia (0-6 años)

En: *Revista de Educación*, Ministerio de Educación, Política Social y Deporte, 347, 2008, pp. 141-153. [[http://www.revistaeducacion.mec.es/re347\\_07.html](http://www.revistaeducacion.mec.es/re347_07.html)]. [ISSN 0034-8082]

## Sección ordinaria

### Investigaciones y estudios

ALCALDE CAMPOS, R.

Los programas de actuación educativa orientados al alumnado de origen extranjero: ¿modelos de atención a la diversidad cultural o a la igualdad educativa?

En: *Revista de Educación*, Ministerio de Educación, Política Social y Deporte, 345, 2008, pp. 207-228. [[http://www.revistaeducacion.mec.es/re345\\_09.html](http://www.revistaeducacion.mec.es/re345_09.html)]. [ISSN 0034-8082]

ÁLVAREZ ROJO, V., GARCÍA JIMÉNEZ, E., FLORES ALÉS, V. Y CORREA MANFREDIX, J.

Aproximación a la cultura de un centro universitario

En: *Revista de Educación*, Ministerio de Educación, Política Social y Deporte, 346, 2008, pp. 139-166. [[http://www.revistaeducacion.mec.es/re346\\_05.html](http://www.revistaeducacion.mec.es/re346_05.html)]. [ISSN 0034-8082]

ANAYA MOIX, V. Y LLOBERA, M.

El tratamiento de la comunicación oral y escrita en los currículos franceses de la Enseñanza del Español como Lengua Extranjera (ELE)

En: *Revista de Educación*, Ministerio de Educación, Política Social y Deporte, 345, 2008, pp. 427-443. [[http://www.revistaeducacion.mec.es/re345\\_18.html](http://www.revistaeducacion.mec.es/re345_18.html)]. [ISSN 0034-8082]

BOSCO PANIAGUA, A., LARRAÍN PFINGSTHORN, V. Y SANCHO GIL, J. M<sup>a</sup>, HERNÁNDEZ Y HERNÁNDEZ, F.

School +: un proyecto europeo para repensar la Enseñanza Secundaria

En: *Revista de Educación*, Ministerio de Educación, Política Social y Deporte, 347, 2008, pp. 157-180. [[http://www.revistaeducacion.mec.es/re347\\_08.html](http://www.revistaeducacion.mec.es/re347_08.html)]. [ISSN 0034-8082]

BRETONES ROMÁN, A.

Participación del alumnado de Educación Superior en su evaluación

En: *Revista de Educación*, Ministerio de Educación, Política Social y Deporte, 347, 2008, pp. 181-202. [[http://www.revistaeducacion.mec.es/re347\\_09.html](http://www.revistaeducacion.mec.es/re347_09.html)]. [ISSN 0034-8082]

CANTÓN MAYO, I. Y ARIAS GAGO, A. R.

La dirección y el liderazgo: aceptación, conflicto y calidad

En: *Revista de Educación*, Ministerio de Educación, Política Social y Deporte, 345, 2008, pp. 229-254. [[http://www.revistaeducacion.mec.es/re345\\_10.html](http://www.revistaeducacion.mec.es/re345_10.html)]. [ISSN 0034-8082]

CASTAÑO, E., GALLÓN, S., GÓMEZ, K. Y VÁSQUEZ, J.

Análisis de los factores asociados a la deserción estudiantil en la Educación Superior: un estudio de caso

En: *Revista de Educación*, Ministerio de Educación, Política Social y Deporte, 345, 2008, pp. 255-280. [[http://www.revistaeducacion.mec.es/re345\\_11.html](http://www.revistaeducacion.mec.es/re345_11.html)]. [ISSN 0034-8082]

CASTEJÓN FERNÁNDEZ, L., GONZÁLEZ-PUMARIEGA SOLÍS, S., NÚÑEZ PÉREZ, J. C. Y GONZÁLEZ-PIENDA GARCÍA, J.A.

La escuela inclusiva ante los alumnos tartamudos: actitudes y emociones en la relación educativa

En: *Revista de Educación*, Ministerio de Educación, Política Social y Deporte, 345, 2008, pp. 281-299. [[http://www.revistaeducacion.mec.es/re345\\_12.html](http://www.revistaeducacion.mec.es/re345_12.html)]. [ISSN 0034-8082]

CECCHINI, J.A., FERNÁNDEZ LOSA, J., GONZÁLEZ, C. Y ARRUIZA, J.A.

Repercusiones del Programa Delfos de educación en valores a través del deporte en jóvenes escolares

En: *Revista de Educación*, Ministerio de Educación, Política Social y Deporte, 346, 2008, pp. 167-186. [[http://www.revistaeducacion.mec.es/re346\\_06.html](http://www.revistaeducacion.mec.es/re346_06.html)]. [ISSN 0034-8082]

COLÁS BRAVO, P. Y JIMÉNEZ CORTÉS, R.

Evaluación del impacto de la formación (online) en TIC en el profesorado. Una perspectiva sociocultural

En: *Revista de Educación*, Ministerio de Educación, Política Social y Deporte, 346, 2008, pp. 187-215. [[http://www.revistaeducacion.mec.es/re346\\_07.html](http://www.revistaeducacion.mec.es/re346_07.html)]. [ISSN 0034-8082]

ESSOMBA GELABERT, M.A. Y GONZÁLEZ-PIENDA GARCÍA, A.

Procesos y dinámicas de construcción identitaria entre adolescentes inmigrados. Hacia una política educativa y cultural del reconocimiento

En: *Revista de Educación*, Ministerio de Educación, Política Social y Deporte, 346, 2008, pp. 217-243. [[http://www.revistaeducacion.mec.es/re346\\_08.html](http://www.revistaeducacion.mec.es/re346_08.html)]. [ISSN 0034-8082]

FERNÁNDEZ ANTELO, I. Y CUADRADO GORDILLO, I.

Adaptaciones cognitivo-lingüísticas del profesor a los comportamientos comunicativos de los alumnos

En: *Revista de Educación*, Ministerio de Educación, Política Social y Deporte, 345, 2008, pp. 301-328. [[http://www.revistaeducacion.mec.es/re345\\_13.html](http://www.revistaeducacion.mec.es/re345_13.html)]. [ISSN 0034-8082]

FERNÁNDEZ SORIA, J. M.

Derecho a la educación y libertad de enseñanza en la reforma educativa (LOE)

En: *Revista de Educación*, Ministerio de Educación, Política Social y Deporte, 346, 2008, pp. 245-265. [[http://www.revistaeducacion.mec.es/re346\\_09.html](http://www.revistaeducacion.mec.es/re346_09.html)]. [ISSN 0034-8082]

FLORES, J., SALGUERO, A. Y MÁRQUEZ, S.

Relación de género, curso y tipo de colegio con el clima motivacional percibido en la Educación Física escolar en estudiantes colombianos

En: *Revista de Educación*, Ministerio de Educación, Política Social y Deporte, 347, 2008, pp. 203-227. [[http://www.revistaeducacion.mec.es/re347\\_10.html](http://www.revistaeducacion.mec.es/re347_10.html)]. [ISSN 0034-8082]

GALLEGO ORTEGA, J. L.

La planificación de la expresión escrita por alumnos con retraso mental

En: *Revista de Educación*, Ministerio de Educación, Política Social y Deporte, 346, 2008, pp. 267-290. [[http://www.revistaeducacion.mec.es/re346\\_10.html](http://www.revistaeducacion.mec.es/re346_10.html)]. [ISSN 0034-8082]

GIL FLORES, J.

Consumo de alcohol entre estudiantes de enseñanzas secundarias. Factores de riesgo y factores de protección

En: *Revista de Educación*, Ministerio de Educación, Política Social y Deporte, 346, 2008, pp. 291-313. [[http://www.revistaeducacion.mec.es/re346\\_11.html](http://www.revistaeducacion.mec.es/re346_11.html)]. [ISSN 0034-8082]

GUERRERO BARONA, E. Y RUBIO JIMÉNEZ, J. C.

Fuentes de estrés y síndrome de Burnout en orientaciones de Institutos de Enseñanza Secundaria

En: *Revista de Educación*, Ministerio de Educación, Política Social y Deporte, 347, 2008, pp. 229-254. [[http://www.revistaeducacion.mec.es/re347\\_11.html](http://www.revistaeducacion.mec.es/re347_11.html)]. [ISSN 0034-8082]

KOLLER, S.

Aquí es fácil perder tiempo: estudio y formación en la experiencia de migrantes ecuatorianos

En: *Revista de Educación*, Ministerio de Educación, Política Social y Deporte, 346, 2008, pp. 315-334. [[http://www.revistaeducacion.mec.es/re346\\_12.html](http://www.revistaeducacion.mec.es/re346_12.html)]. [ISSN 0034-8082]

MARTÍNEZ USARRALDE, M<sup>a</sup>. J.

Una mirada comparada al contexto internacional: revisión de la realidad del mediador cultural en algunos países. El caso de Canadá, Italia, Suecia y Portugal

En: *Revista de Educación*, Ministerio de Educación, Política Social y Deporte, 345, 2008, pp. 353-376. [[http://www.revistaeducacion.mec.es/re345\\_15.html](http://www.revistaeducacion.mec.es/re345_15.html)]. [ISSN 0034-8082]

MATEOS, M., MARTÍN, E., PECHARROMÁN, A., LUNA, M<sup>a</sup> E CUEVAS, I.

Estudio sobre la percepción de los estudiantes de Psicología de la tareas de lectura y escritura que realizan para aprender

En: *Revista de Educación*, Ministerio de Educación, Política Social y Deporte, 347, 2008, pp. 255-274. [[http://www.revistaeducacion.mec.es/re347\\_12.html](http://www.revistaeducacion.mec.es/re347_12.html)]. [ISSN 0034-8082]

MOLINA RUIZ, E., IRANZO GARCÍA, P., LÓPEZ LÓPEZ, M<sup>a</sup> C. Y MOLINA MERLO, M<sup>a</sup> A.  
Procedimientos de análisis, evaluación y mejora de la formación práctica  
En: *Revista de Educación*, Ministerio de Educación, Política Social y Deporte, 346, 2008, pp. 335-361. [[http://www.revistaeducacion.mec.es/re346\\_13.html](http://www.revistaeducacion.mec.es/re346_13.html)]. [ISSN 0034-8082]

MONTERO ALCAIDE, A.  
El acceso a la dirección en el sistema educativo español. Dificultades para la definición de un modelo  
En: *Revista de Educación*, Ministerio de Educación, Política Social y Deporte, 347, 2008, pp. 275-298. [[http://www.revistaeducacion.mec.es/re347\\_13.html](http://www.revistaeducacion.mec.es/re347_13.html)]. [ISSN 0034-8082]

MORENO MURCIA, J.A., GONZÁLEZ-CUTRE COLL, D. Y SICILIA, Á.  
Metas de logro 2 x 2 en estudiantes españoles de Educación Física  
En: *Revista de Educación*, Ministerio de Educación, Política Social y Deporte, 347, 2008, pp. 299-317. [[http://www.revistaeducacion.mec.es/re347\\_14.html](http://www.revistaeducacion.mec.es/re347_14.html)]. [ISSN 0034-8082]

NAVARRO MONTAÑO, M. J.  
La dirección escolar ante el reto de la diversidad  
En: *Revista de Educación*, Ministerio de Educación, Política Social y Deporte, 347, 2008, pp. 319-341. [[http://www.revistaeducacion.mec.es/re347\\_15.html](http://www.revistaeducacion.mec.es/re347_15.html)]. [ISSN 0034-8082]

PÉREZ GARCÍA, M<sup>a</sup> P.  
Competencias adquiridas por los futuros docentes desde la formación inicial  
En: *Revista de Educación*, Ministerio de Educación, Política Social y Deporte, 347, 2008, pp. 343-367. [[http://www.revistaeducacion.mec.es/re347\\_16.html](http://www.revistaeducacion.mec.es/re347_16.html)]. [ISSN 0034-8082]

PINEDA HERRERO, P., MORENO ANDRÉS, M<sup>a</sup> V Y BELVIS PONS, E.  
La movilidad de los universitarios en España: estudio sobre la participación en los programas Erasmus y Sicue  
En: *Revista de Educación*, Ministerio de Educación, Política Social y Deporte, 346, 2008, pp. 363-399. [[http://www.revistaeducacion.mec.es/re346\\_14.html](http://www.revistaeducacion.mec.es/re346_14.html)]. [ISSN 0034-8082]

ROJAS PERNIA, S.  
La «voz» de las personas con discapacidad intelectual en investigación educativa. Repensando las prácticas de investigación  
En: *Revista de Educación*, Ministerio de Educación, Política Social y Deporte, 345, 2008, pp. 377-398. [[http://www.revistaeducacion.mec.es/re345\\_16.html](http://www.revistaeducacion.mec.es/re345_16.html)]. [ISSN 0034-8082]

SÁNCHEZ GARCÍA, M<sup>a</sup>. F. (COORD.), GUILLAMÓN FERNÁNDEZ, J. R. (COORD.), FERRER SAMA, P., VILLALBA VÍLCHEZ, E., MARTÍN CUADRADO, A. M. Y PÉREZ GONZÁLEZ, J. C.  
Situación actual de los servicios de orientación universitaria: estudio descriptivo

En: *Revista de Educación*, Ministerio de Educación, Política Social y Deporte, 345, 2008, pp. 329-352. [[http://www.revistaeducacion.mec.es/re345\\_14.html](http://www.revistaeducacion.mec.es/re345_14.html)]. [ISSN 0034-8082]

TORREGO SEIJO, J. C. Y GALÁN GONZÁLEZ, A.

Investigación evaluativa sobre el programa de mediación de conflictos en centros escolares  
En: *Revista de Educación*, Ministerio de Educación, Política Social y Deporte, 347, 2008, pp. 369-394. [[http://www.revistaeducacion.mec.es/re347\\_17.html](http://www.revistaeducacion.mec.es/re347_17.html)]. [ISSN 0034-8082]

VIEIRA ALLER, M<sup>a</sup> J.

Criterios para la evaluación del sistema de apoyo y orientación al estudiante universitario: revisión y propuesta.

En: *Revista de Educación*, Ministerio de Educación, Política Social y Deporte, 345, 2008, pp. 399-423. [[http://www.revistaeducacion.mec.es/re345\\_17.html](http://www.revistaeducacion.mec.es/re345_17.html)]. [ISSN 0034-8082]

VILA MENDIBURU, J. I.

Lengua familiar y conocimiento de la lengua escolar en Cataluña al finalizar la Educación Infantil

En: *Revista de Educación*, Ministerio de Educación, Política Social y Deporte, 346, 2008, pp. 401-424. [[http://www.revistaeducacion.mec.es/re346\\_15.html](http://www.revistaeducacion.mec.es/re346_15.html)]. [ISSN 0034-8082]

## **Ensayos e informes**

BRUNO-JOFRE, R. Y JOVER OLMEDA, G.

Los estudios de formación docente y pedagógica en Canadá y España: cambios programáticos e institucionales en el escenario de internacionalización de la educación

En: *Revista de Educación*, Ministerio de Educación, Política Social y Deporte, 347, 2008, pp. 397-417. [[http://www.revistaeducacion.mec.es/re347\\_18.html](http://www.revistaeducacion.mec.es/re347_18.html)]. [ISSN 0034-8082]

DEL POZO ORTIZ, A.

La necesaria reforma de la función pública docente recogida en la LOE

En: *Revista de Educación*, Ministerio de Educación, Política Social y Deporte, 345, 2008, pp. 457-479. [[http://www.revistaeducacion.mec.es/re345\\_20.html](http://www.revistaeducacion.mec.es/re345_20.html)]. [ISSN 0034-8082]

GRANADOS ROMERO, J. M<sup>a</sup>.

La alfabetización tecnológica y el acceso a las TIC en la dialéctica inclusión/exclusión. El caso de la minoría gitana

En: *Revista de Educación*, Ministerio de Educación, Política Social y Deporte, 346, 2008, pp. 427-454. [[http://www.revistaeducacion.mec.es/re346\\_16.html](http://www.revistaeducacion.mec.es/re346_16.html)]. [ISSN 0034-8082]

HALBACH, A.

Una metodología para la enseñanza bilingüe en la etapa de Primaria

En: Revista de Educación, Ministerio de Educación, Política Social y Deporte, 346, 2008, pp. 455-466. [[http://www.revistaeducacion.mec.es/re346\\_17.html](http://www.revistaeducacion.mec.es/re346_17.html)]. [ISSN 0034-8082]

MOREU CALVO, A. C.

Psicología y Pedagogía. Consideraciones en torno a la herencia de Wilhelm Wundt

En: Revista de Educación, Ministerio de Educación, Política Social y Deporte, 345, 2008, pp. 443-456. [[http://www.revistaeducacion.mec.es/re345\\_19.html](http://www.revistaeducacion.mec.es/re345_19.html)]. [ISSN 0034-8082]

SERDIO SÁNCHEZ, C.

La educación en la vejez: fundamentos y retos de futuro

En: Revista de Educación, Ministerio de Educación, Política Social y Deporte, 346, 2008, pp. 467-483. [[http://www.revistaeducacion.mec.es/re346\\_18.html](http://www.revistaeducacion.mec.es/re346_18.html)]. [ISSN 0034-8082]

UNCETA SATRÚSTEGUI, A.

Cambios sociales y educación. Notas para el debate

En: Revista de Educación, Ministerio de Educación, Política Social y Deporte, 347, 2008, pp. 419-432. [[http://www.revistaeducacion.mec.es/re347\\_19.html](http://www.revistaeducacion.mec.es/re347_19.html)]. [ISSN 0034-8082]

### **Experiencias educativas (innovación)**

PÉREZ PUEYO, A., TABERNERO SÁNCHEZ, B., LÓPEZ PASTOR, V. M., UREÑA ORTÍN, N., RUIZ LARA, E., CAPLLOCH BUJOSA, M., GONZÁLEZ FERNÁNDEZ, N. Y CASTEJÓN OLIVA, F. J.

Evaluación formativa y compartida en la docencia universitaria y el Espacio Europeo de Educación Superior: cuestiones clave para su puesta en práctica

En: Revista de Educación, Ministerio de Educación, Política Social y Deporte, 347, 2008, pp. 435-451. [[http://www.revistaeducacion.mec.es/re347\\_20.html](http://www.revistaeducacion.mec.es/re347_20.html)]. [ISSN 0034-8082]

SOUTO IGLESIAS, A. Y BRAVO TRINIDAD, J. L.

Implementación European Credit Transfer System en un curso de Programación en Ingeniería

En: Revista de Educación, Ministerio de Educación, Política Social y Deporte, 346, 2008, pp. 487-511. [[http://www.revistaeducacion.mec.es/re346\\_19.html](http://www.revistaeducacion.mec.es/re346_19.html)]. [ISSN 0034-8082]

VENEGAS MEDINA, M.

I Jornadas sobre Coeducación, Igualdad y Buenos Tratos en el colegio San José de Cartuja. Una experiencia coeducativa para la ciudadanía

En: Revista de Educación, Ministerio de Educación, Política Social y Deporte, 345, 2008, pp. 483-497. [[http://www.revistaeducacion.mec.es/re345\\_21.html](http://www.revistaeducacion.mec.es/re345_21.html)]. [ISSN 0034-8082]

**Número Extraordinario 2008. Tiempos de cambio universitario en Europa  
(Coordinador: Francisco Michavila).**

MICHAVILA, F.

**Presentación**

En: Revista de Educación, Ministerio de Educación, Política Social y Deporte, 2008, extraordinario, pp. 15-19. [[http://www.revistaeducacion.mec.es/re2008\\_01.htm](http://www.revistaeducacion.mec.es/re2008_01.htm)]. [ISSN 0034-8082]

VÁZQUEZ GARCÍA, J.A.

**La organización de las enseñanzas de grado y postgrado**

En: Revista de Educación, Ministerio de Educación, Política Social y Deporte, 2008, extraordinario, pp. 23-39. [[http://www.revistaeducacion.mec.es/re2008\\_02.htm](http://www.revistaeducacion.mec.es/re2008_02.htm)]. [ISSN 0034-8082]

GINÉS MORA, J. Y CEGES-LMPF

**El «éxito laboral» de los jóvenes graduados universitarios europeos**

En: Revista de Educación, Ministerio de Educación, Política Social y Deporte, 2008, extraordinario, pp. 41-58. [[http://www.revistaeducacion.mec.es/re2008\\_03.htm](http://www.revistaeducacion.mec.es/re2008_03.htm)]. [ISSN 0034-8082]

HUBER, G. L.

**Aprendizaje activo y metodologías educativas**

En: Revista de Educación, Ministerio de Educación, Política Social y Deporte, 2008, extraordinario, pp. 59-81. [[http://www.revistaeducacion.mec.es/re2008\\_04.htm](http://www.revistaeducacion.mec.es/re2008_04.htm)]. [ISSN 0034-8082]

MICHAVILA, F. Y PAREJO, J. L.

**Políticas de participación estudiantil en el proceso de Bolonia**

En: Revista de Educación, Ministerio de Educación, Política Social y Deporte, 2008, extraordinario, pp. 85-118. [[http://www.revistaeducacion.mec.es/re2008\\_05.htm](http://www.revistaeducacion.mec.es/re2008_05.htm)]. [ISSN 0034-8082]

BRUNNER, J. J.

**El proceso de Bolonia en el horizonte latinoamericano: límites y posibilidades**

En: Revista de Educación, Ministerio de Educación, Política Social y Deporte, 2008, extraordinario, pp. 119-145. [[http://www.revistaeducacion.mec.es/re2008\\_06.htm](http://www.revistaeducacion.mec.es/re2008_06.htm)]. [ISSN 0034-8082]

DOCAMPO, D.

**Rankings internacionales y calidad de los sistemas universitarios**

En: Revista de Educación, Ministerio de Educación, Política Social y Deporte, 2008, extraordinario, pp. 149-176. [[http://www.revistaeducacion.mec.es/re2008\\_07.htm](http://www.revistaeducacion.mec.es/re2008_07.htm)]. [ISSN 0034-8082]

VILLA SÁNCHEZ, A.

**La excelencia docente**

En: Revista de Educación, Ministerio de Educación, Política Social y Deporte, 2008, extraordinario, pp. 177-212. [[http://www.revistaeducacion.mec.es/re2008\\_08.htm](http://www.revistaeducacion.mec.es/re2008_08.htm)]. [ISSN 0034-8082]

MARTÍNEZ MARTÍN, M. Y VIADER JUNYENT, M.

Reflexiones sobre aprendizaje y docencia en el actual contexto universitario. La promoción de equipos docentes

En: Revista de Educación, Ministerio de Educación, Política Social y Deporte, 2008, extraordinario, pp. 213-234. [[http://www.revistaeducacion.mec.es/re2008\\_09.htm](http://www.revistaeducacion.mec.es/re2008_09.htm)]. [ISSN 0034-8082]

MICHAVILA, F. Y ZAMORANO, S.

Panorama de sistemas de Garantía de Calidad en Europa

En: Revista de Educación, Ministerio de Educación, Política Social y Deporte, 2008, extraordinario, pp. 235-263. [[http://www.revistaeducacion.mec.es/re2008\\_10.htm](http://www.revistaeducacion.mec.es/re2008_10.htm)]. [ISSN 0034-8082]

GARCÍA-BENAU, M<sup>a</sup> A.

Algunos aspectos sobre la racionalidad en la financiación universitaria

En: Revista de Educación, Ministerio de Educación, Política Social y Deporte, 2008, extraordinario, pp. 267-284. [[http://www.revistaeducacion.mec.es/re2008\\_11.htm](http://www.revistaeducacion.mec.es/re2008_11.htm)]. [ISSN 0034-8082]

HAUG, G.

La legislación europea y las legislaciones nacionales

En: Revista de Educación, Ministerio de Educación, Política Social y Deporte, 2008, extraordinario, pp. 285-305. [[http://www.revistaeducacion.mec.es/re2008\\_12.htm](http://www.revistaeducacion.mec.es/re2008_12.htm)]. [ISSN 0034-8082]

## **Reseñas bibliográficas del 2008**

ABDELILAH-BAUER, B. (2007). *El desafío del bilingüismo*. Madrid: Morata. 130 pp. ISBN: 978-84-7112-523-1. (347, Marco Antonio García Cortés).

AGUADED GÓMEZ, J. Y FONSECA MORA, C. (2007). *Enseñar en la Universidad: Experiencias y propuestas para la docencia universitaria*. La Coruña, Netbiblo. 308 pp. ISBN: 978-84-9745-081-2. (Número Extraordinario 2008, Francisco Pavón Rabasco)

AMBRÓS, A. Y BREU, R. (2007). *Cine y educación. El cine en el aula de primaria y secundaria*. Barcelona: Graó. 233 pp. ISBN: 978-84-7827-496-3. (347, Octavio Morales Castro).

ARIÑO VILLARROYA, A. (Dir.). (2007). *El oficio de estudiar en la Universidad: compromisos flexibles*. Valencia, PUV. 260 pp. ISBN: 978-84-370-7033-9. (Número Extraordinario 2008, Guillermo Quintás).

BARKLEY, E., CROSS, K. P. Y HOWELL MAJOR, C. (2007). *Técnicas de aprendizaje colaborativo*. Madrid: MEC-CIDE / Morata. 236 pp. ISBN: 978-84-7112-522-4. (Número Extraordinario 2008, Ana María Madrigal Martínez).

BEANE, A. L. (2006). *Bullying. Aulas libres de acoso*. Barcelona: Graó. 245 pp. ISBN:978-84-7827-439-0 (346, Blanca Sotos Vidal).

BERNARDO CARRASCO, J., JAVALOYES SOTO, J. J. Y CALDERERO HERNÁNDEZ, J. F. (2007). *Cómo personalizar la educación. Una solución de futuro*. Madrid: Narcea. 248 pp. ISBN: 978-84-277-1559-2. (Número Extraordinario 2008, Santiago Hernández Abad).

CASTELLÓ, M. ET AL. (2007). *Escribir y comunicarse en contextos científicos y académicos. Conocimientos y estrategias*. Barcelona: Graó. 224 pp. ISBN:84-7827-477-2 (346, Mar Mateos Sanz).

CASTELLS, P. (2007). *Víctimas y matones. Claves para afrontar la violencia en niños y jóvenes*. Barcelona: Ediciones CEAC. 145 pp. ISBN: 978-84-329-1779-0. (345, Lidia Ayllón Romero).

CELORIO, G. Y LÓPEZ DE MUNAÍN, A. (COORDS.) (2007). *Diccionario de Educación para el Desarrollo*. Bilbao: Hegoa. 284 pp. ISBN: 978-84-89916-02-9. (Número Extraordinario 2008, Fernando Castro Vega).

COBO, R. (ed.) (2008). *Educación en la ciudadanía. Perspectivas feministas*. Madrid: Catarata. 192 pp. ISBN: 978-84-8319-346-4. (Número Extraordinario 2008, Ana Franco).

EURYDICE (2007). *La enseñanza de las ciencias en los centros escolares de Europa. Políticas e investigación*. Madrid: MEC-CIDE. 94 pp. ISBN: 92-79-03558-4 (346, Ana María Madrigal Martínez).

FEITO, R. (2006). *Otra escuela es posible*. Madrid: Siglo XXI de España Editores, S. A. 244 pp. ISBN: 978-84-323-1275-5. (345, Cristina Tiernes Cruz).

FERNÁNDEZ, M., SOUTO, X.M. Y RODRÍGUEZ, R. (2007). *La sociedad del conocimiento: democracia y cultura*. Madrid: Secretaría General Técnica del MEC; FIES (Fundación de Investigaciones Educativas y Sindicales) y Barcelona: Ediciones Octaedro. 149 pp. ISBN: 978-84-8063-715-3. (Número Extraordinario 2008, Cristina Tiernes Cruz).

GARCÍA ARETIO, L. (COORD.), RUIZ CORBELLA, M. Y DOMÍNGUEZ FIGAREDO, D. (2007). *De la educación a distancia a la educación virtual*. Barcelona: Ariel. 303 pp. ISBN: 978-84-3442-666-5 (346, Miriam García Blanco).

GARCÍA GUERRA, M. (2007). *La razón mestiza. Agenda intercultural*. Madrid: Centro de Investigación y Documentación (CIDE). 418 pp. ISBN: 978-84-369-4428-0 (346, Elena Michelena García).

GEIS, A. Y LOGÁS, J. (COORDS.) (2006). *Dirigir la escuela 0-3*. Barcelona: Graó. 213 pp. ISBN: 84-7827-4969-3. (345, Lorena González-Piñero Doblas).

GIR (GRUPO DE INVESTIGACIÓN RECONOCIDO) DE LA UNIVERSIDAD DE SALAMANCA (2007). *Convergencias y divergencias de la formación inicial práctica del profesorado en la Europa Mediterránea*. Salamanca: Grupo Editorial Universitario. ISBN 84-8491-816-5 (346, María Jesús Martínez Usarralde).

HIGUERAS GARCÍA, M. (2006). *Estudio de las colocaciones léxicas y su enseñanza en español como lengua extranjera*. Málaga: ASELE. 155 pp. ISBN: 8436943082. (347, Eva Pérez Ramos).

LÓPEZ REILLO, P. (2006). *Una oportunidad para aprender. La dimensión intercultural en la formación del profesorado*. Santa Cruz de Tenerife: Cabildo. 302 pp. ISBN: 978-84-7827-421-5. (345, Miguel Anxo Santos Rego).

LOZANO MARTÍNEZ, J. (2007). *Educación en la diversidad*. Barcelona: Editorial Davinci. 228 pp. ISBN: 978-84-93551 (347, Antonio Bautista García-Vera).

LUIS, A. Y ROMERO, J. (2007) *Escuela para todos, conocimiento académico y geografía escolar en España (1830-1953)*. Santander, Servicio de publicaciones de la Universidad de Cantabria. 468 pp. ISBN: 978-84-8102-439-5. (Número Extraordinario 2008, Ramón López Facal).

MARCHESI, A. (2007). *Sobre el bienestar de los docentes. Competencias, emociones y valores*. Madrid: Alianza editorial. ISBN: 978-84-206-8401-7. (345, Víctor Manuel Rodríguez Muñoz).

MARTÍN ROJO, L. Y MIJARES, L. (eds.) (2007). *Voces del aula: Etnografías de la escuela multilingüe*. Madrid: MEC/CIDE. 254 pp. ISBN: 978-84-369-4429-7. (347, David Poveda).

MARTÍNEZ DE BRINGAS, A. (dir.) (2006). *Teoría y práctica de la educación en derechos humanos*. San Sebastián: Departamento para los Derechos Humanos, el Empleo y la Inserción Social de la Diputación Foral de Navarra. 219 pp. ISBN: 84-96310-98-1. (347, Lorena González-Piñero Doblas).

MARTÍNEZ GONZÁLEZ R. A. (dir.) (2007). *La investigación en la práctica educativa: Guía metodológica de investigación para el diagnóstico y evaluación en los centros docentes*. Madrid: Centro de Investigación y Documentación Educativa (CIDE). Secretaría general técnica. Subdirección general de información y publicaciones. 129 pp. ISBN: 978-84-369-4440-2. (Número Extraordinario 2008, Rosa A. Huertas Mora).

MIJARES MOLINA, L. (2006). *Aprendiendo a ser marroquíes. Inmigración, diversidad lingüística y escuela*. Madrid: Ediciones del Oriente y del Mediterráneo. Colección «Encuentros». Serie TEIM 3. 406 pp. ISBN: 978-84-96327-38-2. (345, Concha Moreno García).

MINISTERIO DE EDUCACIÓN Y CIENCIA (2007). *Calidad para todos. Premio Marta Mata 2007*. Madrid: Secretaría General Técnica. Subdirección General de Información y Publicaciones. 365 pp. ISBN: 978-84-7827-421-5. (345, Elena Vázquez Aguilar).

ORTEGA, J. A. Y CHACÓN, A. (Coords.) (2007). *Nuevas tecnologías para la educación en la era digital*. Madrid: Editorial Pirámide. 406 pp. ISBN: 978-4-368-2086-7 (346, José Luis Villena Higuera).

PÉREZ GÓMEZ, A., MARTÍNEZ, M., TEY, A., ESSOMBA, M. A. Y GONZÁLEZ GONZÁLEZ, M. T. (2007). *Profesorado y otros profesionales de la educación. Alternativas para un sistema escolar democrático*. Madrid: Ediciones Octaedro, 152 pp. ISBN: 978-84-8063-867-8 (346, Gala-Emma Peñalba Esteban).

RAMOS MÉNDEZ, C. (2007). *El pensamiento de los aprendientes en torno a cómo se aprende una lengua: dimensiones individuales y culturales*. Madrid: ASELE. 129 pp. ISBN 978-84-3694-460-0 (346, Óscar Ruiz Maroto).

RAMOS ZAMORA, S. (2006). *La represión del magisterio. Castilla-La Mancha, 1936-1945*. Ciudad Real: Almud ediciones de Castilla-La Mancha. 360 pp. ISBN: 978-84-9341-404-7 (346, Miryam Carreño).

REBOLLO, M. A. (2005). *Género e interculturalidad: educar para la igualdad*. Madrid: La Muralla. ISBN: 978-84-7133-763-4. (345, Concepción Mimbreno Mallado).

RUDDUCK, J. Y FLUTTER, J. (2007). *Cómo mejorar tu centro escolar dando la voz al alumnado*. Madrid: Morata. 171 pp. ISBN: 978-84-7112-518-7. (347, Cristina Tienes Cruz).

SÁNCHEZ BLANCO, C. (2006). *Violencia física y construcción de identidades: Propuesta de reflexión crítica para las escuelas infantiles*. Barcelona: Graó. ISBN: 978-84-7827-421-5. (345, Patricia Digón Regueiro).

SANTANA, L. E. (2007). *Orientación Educativa e Intervención Psicopedagógica*. Madrid: Ediciones Pirámide. 375 pp. ISBN: 978-84-368-2125-3 (346, Cristina Tienes Cruz).

TEJADA FERNÁNDEZ, J. Y JIMÉNEZ MARÍN, V. (Coords.) (2006). *Formación de Formadores. Escenario Aula*. Volumen 1. Madrid: Thomson-Paraninfo. 707 pp. ISBN: 84-9732-519-2. (345, Óscar Mas Torelló).

VERA, J. M., MORA, V. Y LAPEÑA, A. (2006). *Dirección y gestión de centros docentes. Guía práctica para el trabajo diario del equipo directivo*. Barcelona: Graó. 284 pp. ISBN: 84-7827-458-8. (345, Noelia Martínez Mesones).

VV. AA. (2007). *Educación y Diversidad. Education and Diversity. Anuario Internacional de Investigación sobre Discapacidad e Interculturalidad*. Zaragoza: Mira Editores. 305 pp. ISBN: 978-84-8465-223-6 (346, Gerardo Echeita Sarrionandia).

VV. AA. (2007). *Freaconwin: Corpus Canario de Inglés Oral*. Canarias. Consejería de Educación, Cultura y Deportes del Gobierno de Canarias. 212 pp. ISBN: 978-84-690-4238-0. (347, Noelia Martínez Mesones).

WASERMANN, S. (2006). *Jugadores serios en el aula de Primaria: cómo capacitar a los niños mediante experiencias de aprendizaje activo*. Buenos Aires: Amorrortu. 368 pp. ISBN: 950-518-837-4 (347, Consuelo Uceda de Castro).

## Índice de autores

ALCALDE CAMPOS, R.  
ALEGRE CANOSA, M. A.  
ÁLVAREZ ROJO, V.  
ANAYA MOIX, V.  
ARIAS GAGO, A. R.  
ARRUZA, J. A.  
BELVIS PONS, E.  
BLANCO GUIJARRO, R.  
BOSCO PANIAGUA, A.  
BOUACHRA, O.  
BRAVO TRINIDAD, J. L.  
BRETONES ROMÁN, A.  
BRUNNER, J. J.  
BRUNO-JOFRE, R.  
CANTÓN MAYO, I.  
CAPLOCH BUJOSA, M.  
CARRASCO CARPIO, C.  
CASTAÑO, E.  
CASTEJÓN FERNÁNDEZ, L.  
CASTEJÓN OLIVA, F. J.  
CASTELLANO, N.  
CECCHINI, J. A.  
CEGES-LMPF  
COLÁS BRAVO, P.  
COLL, C.  
CORREA MANFREDIX, J.  
CUADRADO GORDILLO, I.  
CUBERO PÉREZ, M.  
CUBERO PÉREZ, R.  
CUEVAS, I.  
DE LA MATA BENÍTEZ, M. L.

DE SIXTE, R.  
DEL POZO ORTIZ, A.  
DOCAMPO, D.  
ESSOMBA GELABERT, M.A.  
FERNÁNDEZ ANTELO, I.  
FERNÁNDEZ ENGUITA, M.  
FERNÁNDEZ LOSA, J.  
FERNÁNDEZ SORIA, J. M.  
FERRER SAMA, P.  
FLORES, J.  
FLORES ALÉS, V.  
FRANZÉ MUDANÓ, A.  
GAETE, J.M.  
GALÁN GONZÁLEZ, A.  
GALLEGO ORTEGA, J. L.  
GALLÓN, S.  
GARCÍA, J. R.  
GARCÍA CASTAÑO, F. J.  
GARCÍA JIMÉNEZ, E.  
GARCÍA-BENAU, M<sup>a</sup> A.  
GARRETA BOCHACA, J  
GIL FLORES, J.  
GINÉS MORA, J.  
GÓMEZ, K.  
GONZÁLEZ, C.  
GONZÁLEZ FERNÁNDEZ, N.  
GONZÁLEZ-CUTRE COLL, D.  
GONZÁLEZ-PIENDA GARCÍA, J.A.  
GONZÁLEZ-PUMARIEGA SOLÍS, S.  
GRANADOS ROMERO, J. M<sup>a</sup>.  
GUERRERO BARONA, E.  
GUIJO BLANCO, V.  
GUILLAMÓN FERNÁNDEZ, J. R.  
HALBACH, A.  
HAUG, G.  
HERNÁNDEZ Y HERNÁNDEZ, F.  
HORNO GOICOECHEA, P.  
HUBER, G. L.  
IGNACIO CARMONA, M. J.  
IRANZO GARCÍA, P.  
JIMÉNEZ CORTÉS, R.  
JOVER OLMEDA, G.

KOLLER, S.  
LARRAÍN PFINGSTHORN, V.  
LEBRERO BAENA, M<sup>a</sup>. P.  
LLOBERA, M.  
LÓPEZ LÓPEZ, M<sup>a</sup>. C.  
LÓPEZ PASTOR, V. M.  
LUNA, M<sup>a</sup>.  
MÁRQUEZ, S.  
MATEOS, M.  
MARTÍN, E.  
MARTÍN CUADRADO, A. M.  
MARTÍNEZ MARTÍN, M.  
MARTÍNEZ USARRALDE, M<sup>a</sup>. J.  
MAURI, T.  
MOLINA MERLO, M<sup>a</sup> A.  
MOLINA RUIZ, E.  
MONTERO ALCAIDE, A.  
MORENO ANDRÉS, M<sup>a</sup>. V.  
MORENO MURCIA, J. A.  
MOREU CALVO, A. C.  
NAVARRO MONTAÑO, M. J.  
NÚÑEZ PÉREZ, J. C.  
ONRUBIA, J.  
PAREJO, J. L.  
PECHARROMÁN, A.  
PÉREZ GARCÍA, M<sup>a</sup> P.  
PÉREZ GONZÁLEZ, J. C.  
PÉREZ PUEYO, A.  
PINEDA HERRERO, P.  
PRADOS GALLARDO, M. M.  
RIESCO SANZ, A.  
ROJAS PERNIA, S.  
ROSALES, J.  
RUBIO GÓMEZ, M.  
RUBIO JIMÉNEZ, J. C.  
RUIZ LARA, E.  
SALGUERO, A.  
SÁNCHEZ, E.  
SÁNCHEZ GARCÍA, M<sup>a</sup>. F.  
SÁNCHEZ SÁNCHEZ, A.  
SANCHO GIL, J. M<sup>a</sup>  
SANTAMARÍA SANTIGOSA, A.

SERDIO SÁNCHEZ, C.  
SICILIA, Á.  
SOUTO IGLESIAS, A.  
TABERNEIRO SÁNCHEZ, B.  
TERRÉN, E.  
TORREGO SEIJO, J. C.  
TORRES SANTOMÉ, J.  
ÚCAR MARTÍNEZ, X.  
UNCETA SATRÚSTEGUI, A.  
UREÑA ORTÍN, N.  
VALCARCE GARCÍA, A.  
VÁSQUEZ, J.  
VÁSQUEZ GARCÍA, J. A.  
VENEGAS MEDINA, M.  
VIADER JUNYENT, M.  
VIEIRA ALLER, M<sup>a</sup> J.  
VILA MENDIBURU, J. I.  
VILLA SÁNCHEZ, A.  
VILLALBA VÍLCHEZ, E.  
ZAMORANO, S.



## **Recensiones y libros recibidos**



# Recensiones y libros recibidos

## Libros reseñados

ALEGRE, M. A. Y SUBIRATS, J. (2007). *Los retos para España en una perspectiva comparada*. Madrid: Centro de Investigaciones Sociológicas, 421 pp. ISBN: 978-84-7476-444-4

COTS, J.M., ARMENGOL, L., ARNÓ, E., IRÚN, M. Y LLURDA, E. (2007). *La conciencia lingüística en la enseñanza de lenguas*. Barcelona: Editorial Graó. 144 pp. ISBN: 978-84-7827-508-3. (Eva Pérez Ramos)

ESCAÑO, J. Y GIL DE LA SERNA, M. (2008). *Cinco hilos para tirar de la motivación y el esfuerzo*. Barcelona: ICE-UB-Horsori. 170 pp. ISBN: 978-849610849-3. (Carmen Robles Casas)

FOUREZ, G. (2008). *Cómo se elabora el conocimiento*. Madrid: Narcea. 183 pp. ISBN: 978-84-277-1572-1. (Irene López Navarro)

HERRÁN DE LA GASCÓN, A. Y PAREDES LABRA, J. (Coords). (2008). *Didáctica General. La práctica de la enseñanza en Educación Infantil, Primaria y Secundaria*. Madrid, McGRAW- HILL/INTER-AMERICANA DE ESPAÑA, S.A.U. 397 pp. ISBN: 978-84-481-6637-3. (Manuela Varilla Valderas)

MARTÍN, X., PUIG, J. M<sup>a</sup>., PADRÓS, M., RUBIO, L. Y TRILLA, J. (2008). *Tutoría. Técnicas, recursos y actividades*. Madrid: Alianza Editorial. 239 pp. ISBN: 978-84-206-8392-8. (Cristina Tiernes Cruz)

MOMINÓ, J.M., SIGALÉS, C. Y MENESES, J. (2008). *La Escuela en la Sociedad red: Internet en la educación Primaria y Secundaria*. Barcelona: Ariel. 257 pp. ISBN: 978-84-344-4275-7. (Fernando Castro Vega)

PALOMARES RUIZ, A. (2007). *Nuevos retos educativos: el modelo docente en el espacio europeo*. Cuenca: Universidad de Castilla La Mancha. 342 pp. ISBN: 978-84-8427-574-9. (Lucía Pérez Inglada)

SÁNCHEZ ARBÓS, M. (2006). *Mi diario*. Zaragoza: Gobierno de Aragón y Caja Inmaculada. 192 pp. ISBN: 978-84-9745-081-2. (Ana Franco)

## Libros recibidos

ARTO, A., PICCINNO, M. Y SERRA, E. (2008). *Comunicar en la educación: Training para la promoción de las habilidades socio-relacionales*. Madrid: CCS.

CAMPS, A. Y MILIAM, M. (COORDS.) (2008). *Miradas y voces: Investigación sobre la educación lingüística y literaria en entornos plurilingües*. Barcelona: Graó.

CUESTA, R. (2007). *Los deberes de la memoria en la educación*. Madrid: Octaedro.

FUNDACIÓN CNSE (2008). *Atención temprana a niñas y niños sordos: Guía para profesionales de los diferentes ámbitos*. Madrid: Autor.

FUNDACIÓN SANTILLANA. (2008). *Las tecnologías de la información y la comunicación (TIC) en la educación: Retos y posibilidades*. Madrid: Santillana.

GÓMEZ YEBRA, A. (2008). *La animación a la lectura entre los jóvenes*. Sevilla: Fundación Ecoem.

LAMET, P. M., ARANZADI, D., RUIZ DE GALARRETA, J.E. ET AL. (2008). *Arrupe y Gárate: dos modelos*. Bilbao: Deusto

LAW, N., PELGRUM, W. Y PLOMP, T. (EDS.). (2008). *Pedagogy and ICT Use in Sschools around the World: Findings from the IEA SITES 2006 Study*. Hong Kong: Comparative Education Research Centre the University of Hong Kong.

MAYER, M. Y TSCHAPKA (EDS.). (2008). *Engaging Routh in sustainable development: Learning and teaching Sustainable Development Lower Secondary Schools*. Bussel: Enviroment and School Iniciatives.

MEPSYD, MINISTERIO DE EDUCACIÓN, POLÍTICA SOCIAL Y DEPORTE (2008). *Bibliotecas escolares: Ideas y buenas y prácticas II*. Madrid: Autor.

MOLINA SAORÍN, J. Y ILLÁN ROMEU, N. (2008). *Educar para la diversidad en la escuela actual: una experiencia práctica de integración curricular*. Sevilla: Editorial MAD

MORA MENDIETA, L. C. (2007). *Concepciones de estudiantes de Licenciatura en Matemáticas sobre números reales*. Bogotá: Universidad Pedagógica Nacional

NAVAL, C. (2008). *Teoría de la Educación: Un análisis epistemológico*. Navarra: Ediciones de la Universidad de Navarra (EUNSA).

RUFFINATO, P. Y SÉIDE, M. (2008). *L'arte di educare nello stile del sistema preventivo: approfondimenti e prospettive*. Roma: Librería Ateneo Salesiano.

**ALEGRE, M. A. Y SUBIRATS, J. (2007). *Los retos para España en una perspectiva comparada*. Madrid: Centro de Investigaciones Sociológicas, 421 pp. ISBN: 978-84-7476-444-4**

Este libro ofrece una visión global sobre el fenómeno de la inmigración vinculado al ámbito educativo escolar. Se trata de una aproximación cualitativa que aúna teoría y práctica sobre el reto que supone la acogida y escolarización de la población escolar inmigrante extranjera.

En su estructura se diferencian seis partes y un primer capítulo introductorio. En esta introducción Miguel Ángel Alegre y Joan Subirats realizan una extensa presentación en la que se reflexiona sobre cuatro debates de interés: la concentración escolar; focalización-universalización de los recursos; aprendizaje de la(s) lengua(s)-educación intercultural y educación-entorno. Todo ello haciendo referencia al resto de capítulos del libro. A la vista de las aportaciones realizadas debe señalarse la extraordinaria labor e interés de sus autores, así como editores de este libro por abrir una vía de debate y reflexión ante la llegada de población extranjera inmigrante a las escuelas.

La primera parte de este libro recoge un capítulo elaborado por Miguel Ángel Alegre, Jordi Collet, Sheila González y Ricard Benito, en él se presenta una versión resumida del informe «La visión comparada del alumno inmigrante. Los casos de Baviera, la comunidad francesa de Bélgica, Inglaterra y Québec» elaborado por el Instituto de Gobierno y Políticas Públicas (IGOP) de la Universidad Autónoma de Barcelona. Aquí se ofrece una serie de datos de tipo descriptivo e interpretativo de las actuaciones que se plantean las autoridades competentes en educación en los países objeto de análisis. Se destaca la evolución y cambios que han ido sufriendo los distintos países en el debate y práctica política referida al tema que nos ocupa. El interés de este capítulo radica en la visión comparada de los mecanismos de gestión de la lengua materna y de acogida, de dinámicas escolares y actuaciones que parten del entorno escolar. Ello demuestra la importante diversidad de modelos de intervención y permite extraer lo mejor de cada uno, con el objeto de reflexionar sobre posibles actuaciones integradoras y normalizadoras.

De la segunda a la sexta parte (capítulos 2-11) se recogen ponencias del seminario internacional «Educando al alumnado inmigrante. Aproximaciones y políticas para la integración educativa y la igualdad de oportunidades para todos» (2006, Barcelona) coordinado por el Instituto de Gobierno y Políticas Públicas y patrocinado por la Fundació Jaime Bofill. En cada una de estas partes se desarrolla respectivamente el caso de Alemania, Francia, Bélgica, Inglaterra y Québec. Al respecto debe destacarse la pluralidad de modelos organizativos que funcionan en Baviera (Alemania) y Québec (Canada).

En la séptima y última parte (capítulos 12-15) se incluye una serie de aportaciones de diversos expertos en el estudio del caso español. Entre las contribuciones de los diversos capítulos de esta parte cabe destacar las realizadas por F. Javier García Castaño, José Fernández Echeverría, María Rubio Gómez y M. Lourdes Soto Páez. Los autores hacen una exhaustiva revisión sobre los conocimientos que la investigación científica en el Estado español ha aportado recientemente sobre la escolarización de la población inmigrante extranjera, sobre la interpretación de las propuestas internacionales de modelos de intervención educativa en término de educación inter o multicultural; así como la normativa que las comunidades autónomas del Estado español han desarrollado recientemente para «acoger» al alumnado extranjero en las aulas y enseñarles la lengua vehicular de la escuela.

La obra supone una notable contribución al conocimiento del panorama, cabe destacar el tratamiento de la temática desde un punto de vista político, ya que tradicionalmente se ha venido poniendo el acento en la cuestiones de carácter únicamente pedagógico. Resulta pues un trabajo muy sugerente, rico en apreciaciones críticas, que aporta una visión comparativa de los diversos casos de estudio no limitándose a su mera exposición. El lector encontrará en esta obra un rico material que le permitirá adquirir un profundo conocimiento en la temática y que le incitará a la reflexión y búsqueda de fórmulas alternativas o mixtas en el contexto de la intervención socioeducativa y escolar.

**Mónica Ortiz Cobo**

**COTS, J.M., ARMENGOL, L., ARNÓ, E., IRÚN, M. LLURDA, E. (2007). *La conciencia lingüística en la enseñanza de lenguas*. Barcelona, Editorial Graó. 144 pp. ISBN: 978-84-7827-508-3.**

Este libro plantea la pregunta de la necesidad de la reflexión sobre la lengua en el aprendizaje de la misma, tanto si esta es materna o un segundo idioma. Y, de este modo, da pautas sobre cómo orientar la educación para fomentar la *conciencia lingüística* en el alumno; hecho que traería como consecuencia una «mejora de sus habilidades comunicativas» (p. 10).

Esta obra parte de la premisa de que «a menudo el lenguaje ha sido estudiado como un cadáver», centrándose solo en la forma, en el análisis de la misma (gramática) y no en el habla, en la actuación comunicativa. Algo de lo que ya hablaban Saussure en su Curso de Lingüística general y Noam Chomsky, el padre de la escuela generativista. Probablemente el impulso de la pragmática en los estudios filológicos impulsó todo este movimiento pro comunicación. Los autores promueven partir siempre de un contexto comunicativo para poder describir mejor la lengua que hablamos o aprendemos.

Para esto se suscita el uso de actividades en el aula que impliquen un análisis lingüístico: a través de textos, diálogos, resolución de problemas, inferencias de textos hablados o escritos... Y también trabajar aspectos como el de la pronunciación (fundamental, porque es un trabajo diario), escritura (dirigida y pautada) y lectura (adecuada a la edad y contexto sociocultural del

alumno). Todos ellos en un marco comunicativo y en un trabajo grupal donde los individuos interactúan y aprenden juntos. Se pretende de algún modo romper la dualidad extrema del panorama educativo español entre forma y comunicación. Es cierto que la importancia de la comunicación y la reflexión de la lengua deriva en comprender lo que hacemos: la estructura interna. Pero no es bueno adoptar solo la postura formal ni solo la comunicativa. «Hay que encontrar maneras de combinar forma y función o lo que es lo mismo: reflexión y comunicación.» (p. 33) Partiendo de la forma (oración, frase, palabra, texto...), podemos deducir las funciones del lenguaje propias a ella en un contexto comunicativo, y toda reflexión debe darse tras el uso social de esa forma, extralingüístico, que hace el hablante de la lengua.

Un buen manual, breve y muy conciso, para profesores de secundaria y bachillerato que reflexionen sobre su labor docente.

**Eva Pérez Ramos**

**ESCAÑO, J., GIL DE LA SERNA, M. (2008). *Cinco bilos para tirar de la motivación y el esfuerzo*. Barcelona: ICE-UB-Horsori. 170 pp. ISBN: 978-849610849-3.**

«Este libro se estructura en capítulos que incluyen una primera parte de exposición de ideas y otra segunda de materiales. En la primera parte, cada uno de los apartados se desarrolla con preguntas que suponen una sucesión de inquietudes y respuestas a los problemas escolares. Los materiales que ofrecemos al final de cada uno de los capítulos tienen como objetivo común servir de recurso práctico».

El libro comienza activando las ideas previas del lector sobre motivación y esfuerzo. Y se cae en la cuenta de que constituyen dos términos de los que casi todos nos quejamos, por su ausencia, y casi nadie interviene a fin de mejorar. Esta circunstancia, inexplicable, lo es más si consideramos que puede apreciarse desde los primeros años de la escolaridad.

Ahora bien, tras una lectura reflexiva de estas primeras páginas, se concluye que la solución es compleja, de entrada por la cantidad de variables y contextos que intervienen. Pero esto, lejos de ser un problema, los autores lo consideran un amplio abanico de posibilidades desde donde intervenir. Dos actividades, aparentemente sencillas, cierran esta introducción a la motivación y el esfuerzo, igualmente desde la reflexión y activación de ideas previas.

A continuación, se presenta el planteamiento de los autores: la motivación y el esfuerzo como capacidades que se desarrollan con la acción educativa (se enseñan y se aprenden), que movilizan y sostienen la actividad intelectual ante cualquier tipo de circunstancia (buena o mala) y que están relacionadas con los buenos resultados académicos. Y para desarrollar estas capacidades, la obra ofrece un repertorio de recursos, a modo de sugerencia de reflexión, que abarca distintos niveles y ámbitos: el sistema escolar, el aula, el propio alumno y la familia. No sin antes desgajar qué actitudes, pensamientos y comportamientos presenta el alumno motivado y ponerlos en relación con determinadas intervenciones de los profesores.

En el capítulo cuarto se plantea el centro educativo como espacio privilegiado para formar la capacidad de motivación y esfuerzo.

Del análisis de las variables fundamentales que caracterizan el sistema escolar, se distinguen cinco grandes variables que constituyen otros tantos ámbitos de intervención o cinco «Hilos para tirar» de la motivación y el esfuerzo.

- Promover *proyectos personales*. De la misma forma que el sistema educativo se estructura en tramos que se han de ir superando para pasar al siguiente, se puede ayudar a tener proyectos asequibles y a organizarse para alcanzar proyectos a más largo plazo.
- Promover el *interés por los contenidos*. Se trata de potenciar aprendizaje significativo, partiendo de los conocimientos previos y tratando de que los nuevos aprendizajes resulten funcionales para los alumnos.
- Promover el *sentimiento de competencia para el aprendizaje*. Especialmente porque cuando los alumnos sienten que no tienen control sobre lo que deben hacer y que están expuestos a un fracaso, se producen sentimientos negativos que desembocan en abandono y evasión.
- Promover la *buena relación con los profesores*. Introduciendo cambios en la organización del centro se articulan tiempos y espacios que facilitan una atención personalizada al alumno.
- Promover la *buena relación con los compañeros*. Tanto de cuestiones puramente escolares como de habilidades sociales y afectivas. Dos propuestas: los alumnos ayudantes de convivencia y los alumnos tutores.

Esta misma estructura se repite en los dos capítulos siguientes, descendiendo al nivel de aula. Se incide en la necesidad de otorgar un papel activo al alumno y de partir del propio estilo docente. Se ofrecen propuestas de intervención para generar reflexión (en profesores y alumnos) y otras referidas a los aspectos metodológicos, concretamente a considerar los espacios y los materiales como recursos que favorecen la motivación y el esfuerzo.

En el séptimo capítulo se analizan las situaciones escolares donde el alumno pone en juego su motivación y esfuerzo. Éstas definen una serie de indicadores de evaluación, útiles para concretar el punto de partida (se dan diversas necesidades educativas) y el objetivo de la intervención (ajustarse a la capacidad de implicación del alumno). Sin olvidar el indispensable traspaso progresivo de control, que va indicando la constancia de los progresos.

El capítulo octavo se ocupa de esa otra estrategia decisiva del aprendizaje: el trabajo del alumno en casa. Es necesario, difícil y genera muchas desigualdades. Los autores proponen convertir el estudio en deberes: el profesor incluye la enseñanza del estudio en sus explicaciones y el alumno materializa su aprendizaje en un cuaderno, que es supervisado.

El libro concluye reflexionando sobre la necesaria colaboración familia-escuela. Nuevamente se recurre a los «hilos para tirar», en este caso, dirigido a los padres. Se ofrecen, además, interesantes ideas para hacer en la vida diaria. Un valioso cuestionario para familias pone el punto final a esta motivadora obra.

**Carmen Robles Casas**

**FOUREZ, G. (2008). *Cómo se elabora el conocimiento*. Madrid: Narcea, 183 pp. ISBN: 978-84-277-1572-1.**

Ocurre en todos los ámbitos científicos sin excepción. También en el de las Ciencias de la Educación. Una vez que acaba la etapa formativa y comienza el ejercicio profesional aparece un énfasis desmesurado en los aspectos metodológicos en pos de una devaluación –o incluso desaparición– de los fundamentos epistemológicos que sustentan dichos procedimientos. Este libro constituye una invitación a no dejarnos llevar por la pereza ni por las actitudes dogmáticas, es decir, a formular la decisiva y molesta pregunta de ¿por qué sabemos lo que sabemos e ignoramos lo que ignoramos? (Lamo de Espinosa, 1994) y, en consecuencia, ¿por qué enseñamos lo que enseñamos?

La respuesta no tiene nada que ver con la casualidad. Tiene que ver con la epistemología. Una rama de la filosofía que permite «estudiar la manera cómo se conoce» (p. 9) y que el autor de este libro ha tratado de conectar con la enseñanza a través de las implicaciones que tienen la una en la otra: «Cuando el docente llega a comprender mejor cómo se construye el saber introduce una nueva dimensión en el tratamiento y la enseñanza de los contenidos disciplinares» (p.10).

La estructura interna de la obra tiene un alto grado de coherencia. Se parte de una introducción general a la epistemología –prólogo– para después presentar una consecución de capítulos de complejidad creciente en los que se dedica un espacio a la explicación de los elementos claves que operan dentro de esta disciplina. Finalmente, llegamos a la parte final de la obra en la que la epistemología (antes tratada en abstracto) se concreta mediante su aplicación al campo de las ciencias de la educación proponiendo una reflexión a sus profesionales acerca de la enseñanza de contenidos científicos.

El enfoque por el que opta el autor del libro es el socio constructivismo y es de agradecer la honestidad de su planteamiento. Ésta le lleva al autor a afirmar desde un principio que el constructivismo es únicamente uno de los puntos de vista posibles –no necesariamente el mejor– para abordar la epistemología. De hecho, lo que hace Fourez al tomar este postulado como punto de partida es precisamente poner en práctica su propio enfoque. El socioconstructivismo afirma que todo conocimiento es situado, es decir, que depende de la posición del sujeto, y que la ciencia es uno de los muchos tipos de conocimientos posibles. Para argumentarlo primero profundizará en la noción de representación como elemento central para entender qué es el conocimiento. Todo saber es una representación que nosotros nos hacemos del mundo. Pero dichas representaciones no son fiel reflejo de la realidad sino que son limitadas y dirigidas a enfocar ciertos elementos de la realidad y no otros. Este hecho contribuye a poner en cuestión las nociones de neutralidad y objetividad tradicionalmente asociadas a toda producción científica. Si repasamos, tal y como se hace en este libro, cómo se ponen a prueba, se aceptan o se rechazan los conocimientos científicos nos daremos cuenta de que éstos se acercan más al devenir histórico que a la intemporalidad. Sin embargo, la corriente positivista continúa siendo la imperante en las comunidades científicas, las cuales trasladan a la sociedad sus descubrimientos como resultado de un conocimiento aséptico y universal. Esta inercia se contagia igualmente a las aulas de las universidades como a las de los colegios, de tal manera que

los maestros reproducen un discurso científico acrítico y exento de valores y cuestionamientos éticos. Pero ¿es correcta esta forma de transmitir conocimientos científicos a los alumnos?, ¿es ésta la fórmula mediante la cual van a «descubrir el mundo» tal y como proponen algunos de los programas escolares? La respuesta de Fourez es tajante. La racionalidad científica no debe ir nunca separada de la ética. Los relatos objetivos acerca de hechos científicos no son el mejor mecanismo pedagógico para la transmisión del conocimiento. Al menos si queremos que los estudiantes entiendan la construcción de los conocimientos como una aventura humana maravillosa que forma parte de la historia.

Irene López Navarro

**HERRÁN DE LA GASCÓN, A. Y PAREDES LABRA, J. (Coords). (2008). *Didáctica General. La práctica de la enseñanza en Educación Infantil, Primaria y Secundaria*. Madrid: McGRAW- HILL/INTERAMERICANA DE ESPAÑA, S.A.U. 397 pp. ISBN: 978-84-481-6637-3.**

Los diferentes autores de esta obra realizan un estupendo acercamiento a la Ciencia de la Didáctica General, para el presente siglo XXI, deteniéndose en sus diferentes problemáticas a la vez ofrecen una visión general y pormenorizada.

Comienzan explicando cómo se puede hacer una buena didáctica, teniendo en cuenta los diferentes dilemas educativos surgidos a través de los años (escolarización vs. educación; aprendizaje superficial vs. aprendizaje significativo,...). Esto debe apoyarse en una teoría de la enseñanza y del currículum, y en cómo ésta afecta a la práctica docente.

Según lo expuesto en la Ley Orgánica 2/2006 de Educación (LOE), los docentes, de acuerdo con su realidad educativa, deben reflexionar para realizar una selección y organización de contenidos que plasmarán en su Programaciones de Aula. A través de los distintos tipos de contenidos, sobre todo actitudes y valores, se favorecerá el desarrollo de las competencias básicas de los alumnos y alumnas y su formación integral. Además para llevar a cabo el proceso de enseñanza-aprendizaje, es necesario utilizar una adecuada metodología, para lo que los autores de la obras muestren diferentes pautas para los diferentes niveles educativos (Educación Infantil, Primaria y Secundaria), procurando fomentar al máximo la creatividad, como factor potenciador del crecimiento personal y de su mejora social.

Por otro lado, la motivación se presenta como un importante factor para el aprendizaje, hecho que recogen los creadores de este libro al tiempo que proponen pautas didácticas que la potencien.

Además, se pone especial interés al presentar al profesor cómo gestor del aula. Se relacionan, entre otras, algunas estrategias para ello: profesor mediador y potenciador del aprendizaje; experto en metodologías; estilos de autoridad y resolución de conflictos.

Otro factor importante en la práctica docente son los recursos didácticos utilizados en el aula, los cuales están sufriendo una importante revolución con la implantación de las

Tecnologías de la Información y la Comunicación. Entre otros, se sugiere el uso de Internet como recurso digital favorecedor de encuentro entre culturas.

Los autores resaltan la importancia de la evaluación de todo el proceso educativo, poniendo especial énfasis en la evaluación de la enseñanza, como elemento de calidad, pues con esta evaluación, se pueden detectar, entre otros, las debilidades y fortalezas de la metodología didáctica empleada.

Pero como no sólo se enseña y aprende lo que viene escrito explícitamente en los distintos documentos curriculares, es necesario tener en cuenta el currículo oculto, y atender a la diversidad presente en nuestras aulas; para lo que proponen realizar diferentes modificaciones curriculares que se adecuen a las necesidades detectadas.

No obstante, aunque todos los factores mencionados hasta hora son necesarios para una buena práctica docente, no son suficientes. Se requiere también una buena formación del profesorado. Este extenso proceso, que no culmina con la formación universitaria inicial, debería prolongarse con una formación permanente. A la vez, para mejorar como profesional de la enseñanza, sería deseable mostrar un carácter investigador e innovador y estar siempre abierto a nuevas experiencias, recursos, métodos, ... es decir, a cambios o modificaciones.

Haciendo una reflexión sobre lo expuesto se observa que todas estas ideas están relacionadas y que es el mismo profesor, desde su práctica cotidiana, el que debe fomentar e implicarse en la mejora del sistema educativo, favoreciendo el enriquecimiento de una escuela inclusiva, una Escuela Para Todos.

Con esta reseña se pretende invitar a educadores, a futuros profesores, a orientadores y a otros profesionales de la educación a realizar una lectura realista y reflexiva de esta obra, en la que se presentan y estudian diferentes aspectos para mejorar y enriquecer la práctica docente.

**Manuela Varilla Valderas**

**MARTÍN, X., PUIG, J. M<sup>a</sup>., PADRÓS, M., RUBIO, L. Y TRILLA, J. (2008). *Tutoría. Técnicas, recursos y actividades*. Madrid: Alianza Editorial. 239 pp. ISBN: 978-84-206-8392-8.**

Con un atractivo color anaranjado se nos presenta esta obra que examina con una mirada profunda la función tutorial ejercida por los profesores de la Enseñanza Secundaria Obligatoria (E.S.O), al mismo tiempo que analiza con una mirada crítica el hecho de que esta labor tutorial no esté plenamente reconocida por la sociedad actual.

Comenzaremos reflexionando acerca de esta *mirada profunda de la labor tutorial* diciendo que, aunque de todos es sabido que los docentes son al mismo tiempo formadores, ya que se encargan de la tarea formativa e instrucción de sus alumnos, estos además, de forma añadida, ejercen una labor tutorial. Los docentes prestan una especial atención a la importancia de educar a los alumnos en valores moralmente trascendentales, para que sean capaces de llevar una vida saludable, y sepan cómo vivir con plena autonomía en las diferentes etapas de su vida.

Una de las contribuciones más reveladoras de este manual es que ofrece a los lectores las pautas para poder elaborar un *Plan de Acción Tutorial (PAT)* a través de una propuesta concreta. Es así como, con un elegante estilo, los autores perfilan un curso académico completo especificando qué se tendría que hacer con el alumnado, con el grupo-clase, con las familias, en el centro, con el equipo educativo y cómo trabajar con los agentes externos.

A medida que avanzamos en la lectura, y con una estilizada pluma, se nos aclara la mejor forma de abordar estas tutorías especificándonos una amplia propuesta de temas para trabajar la educación en valores, cómo intervenir y cómo evaluarlos. Nos proponen además la forma de diseñar y programar las tutorías, es decir, cómo tiene que hacerse la acogida de los alumnos al comienzo de las clases, cómo deben planificarse las entrevistas con familias y los alumnos, y cómo desde las tutorías se ha de fomentar la participación de las familias en el centro, e incluso nos orientan para la resolución de conflictos en las aulas o en el centro.

En uno de sus capítulos centrales se nos presentan las metodologías más adecuadas para emprender esta labor tutorial en un momento y circunstancia concreta de la vida académica de un alumno, encontrándonos así entre sus páginas cuáles serían las formas más idóneas de poner en práctica ciertas técnicas durante la labor tutorial como el role-playing, el diálogo, las habilidades sociales, las técnicas autobiográficas y el uso de recursos tradicionales como la literatura y el periodismo, al mismo tiempo que se nos descubren recursos novedosos que también pueden usarse para tal fin, como pueden ser los audiovisuales, el cine y la fotografía.

Una vez que el profesor-tutor conoce a fondo las diferentes técnicas, será capaz de emplearlas para enseñar valores moralmente relevantes a sus alumnos, como la igualdad de oportunidades para todos, la educación para la paz, la salud, la educación ambiental, trabajar la interculturalidad y los derechos humanos desde las aulas, y para poder abordar con los alumnos y sus familias temas más comprometidos, relacionados con las relaciones afectivo-sexuales de sus hijos. Educando así se consigue una escuela democráticamente implicada con la sociedad que le rodea y de las que consiguen que desde la educación se emanen los valores sociales necesarios para saber convivir.

Finalmente, los autores terminan ofreciendo un amplio elenco de actividades prácticas a desarrollar en las tutorías y así poder fomentar la convivencia y sus normas cívicas, implicar a la familia y transmitir, tanto a los alumnos como a sus familias, la mejor forma de utilizar los recursos que nos ofrece el entorno.

Para terminar, analizaremos esta *mirada crítica de la labor tutorial* que los autores a través de sus páginas nos transmiten. Por ello comenzaremos reflexionando la razón por la cual mientras que en cualquier ámbito empresarial dedicado a otro tipo de actuaciones las responsabilidades ejercidas por sus trabajadores son recompensadas a nivel salarial y con reducciones compensatorias de su jornada laboral, en el ámbito profesional dedicado a la enseñanza, ya sea a nivel público como privado, se siguen presentando este tipo de carencias que consiguen que en España el profesorado esté desmotivado y piense que no les merece la pena ejercer el cargo de responsabilidad que conlleva ser profesor-tutor. Por ello, no es comprensible que actualmente esta amplia labor tutorial ejercida por el profesorado actual, con funciones añadidas, no se vea reconocida y presente deficiencias para quienes la ejercen, como la falta de tiempo, de recursos, de materiales o de una remuneración insuficiente.

Surge por ello una obra como ésta, con una excelente dimensión práctica, que sirve de ayuda con su lectura a todos los profesionales del ámbito educativo que acudan a sus páginas en busca de respuestas, y que por otro lado, propone una apremiante reflexión sobre la falta de medios e incentivos que conlleva el ejercicio de la labor tutorial.

**Cristina Tiernes Cruz**

**MONINÓ, J. M., SIGALÉS, C., MENESES, J. (2008). *La Escuela en la Sociedad red: Internet en la educación Primaria y Secundaria*. Barcelona: Ariel. 257 pp. ISBN: 978-84-344-4275-7.**

Esta investigación, llevada a cabo por los autores Joseph M<sup>a</sup> Mominó, Carles Sigalés y Julio Meneses, analiza en profundidad el desarrollo e implantación de las nuevas tecnologías y sus aplicaciones pedagógicas en la Educación Primaria y Secundaria, así como la descripción de los procesos tecnopedagógicos que se desarrollan en este entorno.

Para abordar cuestiones como éstas, el libro se estructura de la siguiente forma. En el capítulo primero, Introducción, ofrece una mirada teórica y conceptual de la interacción de las Tecnologías de la Información y de la Comunicación (TIC) no sólo en un marco social y educativo emergente, sino que también lleva a cabo un análisis en profundidad, actualizado con datos empíricos del uso, la actuación y el aprovechamiento de las tecnologías que ofrece la red -Internet- en comunidades educativas, dentro y fuera de la escuela, así como los resultados de esta transformación, que supone la inmersión de las nuevas tecnologías en la Educación.

En el capítulo segundo, ofrece una visión panorámica sobre aspectos relacionados con los diseños tecnológicos basados en modelos pedagógicos y con los cambios producidos a raíz de la introducción de las tecnologías de la información, en la mejora, en el grado de desarrollo y en la contribución desde un punto de vista cualitativo, de los sistemas y prácticas pedagógicas en los procesos de enseñanza y aprendizaje.

El tercer capítulo destaca el carácter exploratorio y pionero de este trabajo, el cual parte de una metodología empírica y analítica, que ofrece datos sobre el modo de incorporar las nuevas tecnologías en las escuelas e institutos de Cataluña, a partir de una muestra significativa de centros docentes. Este punto analiza el fenómeno desde dos perspectivas (educativa y tecnológica) con Internet como referencia. Concluye el capítulo con un estudio extensivo que permite conocer cómo se usa Internet en las aulas. También muestra la presencia que tiene la red en los procesos de gestión, organización, dirección y comunicación de los centros.

El siguiente presenta, en forma de síntesis, las principales claves de la situación en la que se encuentra el proceso de integración de Internet en los diversos ámbitos de la actividad de las escuelas e institutos de Cataluña. Estudia la frecuencia, la forma y el papel de las tecnologías de red en el desarrollo de la docencia.

El quinto, analiza en profundidad los aspectos clave, los factores, la frecuencia, la forma, las competencias, y la formación del profesorado en la utilización de las nuevas tecnologías, las

percepciones, las creencias y tipos de prácticas educativas que adopta, los recursos tecnológicos y los apoyos de que dispone, el tipo de centro en que trabaja o la edad, entre otros. En este capítulo se concibe al profesor como el protagonista de la integración de las TIC para la mejora educativa.

En el capítulo sexto el foco prioritario de análisis es la manera en la que los alumnos utilizan la red fuera de la escuela. Además, identifica las consecuencias de la apropiación desigual de las TIC en el ámbito social.

El capítulo séptimo ahonda, desde una perspectiva comunitaria, en el papel de la incorporación de la red como instrumento de colaboración, implicación y participación en la actividad de las escuelas. Este punto expone una concepción socio-tecnológica de los usos de Internet en la vida cotidiana de la comunidad docente, incluyendo alumnos. A través de esta concepción, muestra la contribución específica de las nuevas tecnologías en el desarrollo de las habilidades docentes en diferentes contextos.

En el capítulo octavo y final se presenta, como conclusión, un resumen de las principales aportaciones de este estudio al conocimiento de los procesos de integración de Internet a la educación escolar, y de las implicaciones para la práctica profesional y para el desarrollo de futuras investigaciones. En la segunda parte del capítulo los autores exponen una serie de recomendaciones para elaborar políticas y nuevas exploraciones de apoyo a la docencia.

Conviene destacar que en el anexo del libro se explica de forma detallada la metodología llevada a cabo para realizar el estudio.

**Fernando Castro Vega**

**PALOMARES RUIZ, A. *Nuevos retos educativos: el modelo docente en el espacio europeo*. Cuenca: Universidad de Castilla La Mancha. 342 pp. ISBN: 978-84-8427-574-9.**

Lejos queda ya la Conferencia de Bolonia en la que distintos países europeos elaboraron aquella primera declaración conjunta con el objetivo de armonizar los distintos sistemas educativos. La convergencia europea, no solo supone un cambio en la política educativa y legislación sino incluso un cambio metodológico de los sistemas de enseñanza-aprendizaje.

Ascensión Palomares Ruiz relata las claves y describe del Espacio Europeo con una minuciosa enumeración de los objetivos, cambios y funciones de cada parte implicada en ello; sin olvidar que la educación superior requiere previamente una educación básica previa.

Para tratar este ámbito de organización educativa, la autora comienza explicando el sistema educativo español. Empieza con un breve, pero acertado recorrido por los distintos sistemas educativos de los últimos años, explicando la actual organización descentralizada actual del sistema educativo y la financiación de éste. Finalmente concluye con la exposición de una serie de resultados educativos, para lo que se apoya en distintos indicadores de la educación y en los datos generados por el informe PISA (2003). Obtiene como resultado,

que es imprescindible la mejora de la calidad del sistema educativo Español. La autora, en este análisis crítico sobre la situación educativa en nuestro país, no olvida señalar algunos problemas que están a la orden del día en las aulas; como es la integración del alumno inmigrante o la violencia escolar.

Posteriormente, señala unas propuestas que pueden ayudar a mejorar la calidad del sistema educativo; tanto para docentes, alumnos y administración educativa, como para los propios centros.

El cuarto capítulo analiza los retos de la educación en Europa, partiendo de los problemas actuales de la educación comunes a distintos países y concretando algunas conferencias y declaraciones que se han llevado a cabo los últimos años, con el fin de conseguir mejorar la educación y lograr la convergencia europea. Así señala, por ejemplo, los objetivos de la Conferencia de Bolonia de 1999, y describe al detalle el sistema europeo de transferencia y acumulación de créditos (ECTS); imprescindible para conseguir establecer el EEES.

El profesorado adquiere un papel muy importante a la hora de mejorar la calidad de la educación, y la autora es plenamente consciente de esto. Este penúltimo capítulo describe las distintas alternativas metodológicas y propuestas alternativas, comparándolas con la metodología tradicional, y tratando la figura del profesor no como un transmisor de conocimientos, sino como una persona capaz de guiar a sus alumnos en su propio aprendizaje, sacando provecho de las nuevas tecnologías y de otras metodologías. Así, resume el perfil del docente, detallando las competencias de éstos como son la relevancia de los contenidos que expongan, la planificación del proceso enseñanza-aprendizaje, la comunicación didáctica, el diseño de la metodología y las tecnologías aplicadas a la educación, la función tutorial y finalmente la evaluación democrática. Antes de finalizar el capítulo, enumera una serie de indicadores que podrían ser útiles para medir la calidad de la enseñanza.

En el capítulo sexto estudia el cambio paradigmático que debe producir el docente en el alumnado mediante estrategias y estimulación, con el fin de que éstos obtengan una efectiva inserción social y desarrolle todas sus capacidades para la convivencia y la autorrealización personal, profesional y laboral frente al individualismo y la competitividad de la sociedad actual. Posteriormente describe perfectamente una serie de metodologías activas muy interesantes que pueden introducirse en la docencia universitaria; analiza «el portafolio» apoyándose en un interesante estudio realizado con profesores y estudiantes de distintas universidades, continua con el «aprendizaje colaborativo» señalando como con éste se mejoran algunas deficiencias del conocido «trabajo en grupo», trata también «el aprendizaje basado en problemas», y finalmente «el método del caso».

El libro es seguro de gran interés para estudiantes y profesores universitarios; es de fácil y ágil lectura. Está muy bien estructurado e intercalando una gran cantidad de información con muchas tablas y gráficos ilustrativos que facilitan la comprensión de los distintos conceptos. Si después de leer el libro hubiese alguna duda puede consultarse los anexos finales, en los que está presente toda la legislación relativa a este tema en el momento de su publicación.

**Lucía Inglada Pérez**

**SÁNCHEZ ARBÓS, M. (2006). *Mi diario*. Zaragoza: Gobierno de Aragón y Caja Inmaculada. 192 pp. ISBN: 978-84-9745-081-2.**

Precedido de una presentación de Víctor M. Juan Borroy, un prólogo interesantísimo y entrañable escrito por la pluma de la hija de autora, Elvira Ontañón, y una introducción muy completa debida al ya citado Víctor M. Juan Borroy y a Antonio Viñao Frago, el cuerpo central de *Mi diario* está compuesto por el registro de las notas diarias de María Sánchez Arbós al que se suma, en esta edición, un conjunto de artículos sobre diversos aspectos de la actividad docente.

En *Mi diario*, la autora recoge sus reflexiones del día a día sobre la enseñanza; la labor del maestro y los problemas a los que se enfrenta; la vieja escuela tradicional y la nueva escuela innovadora... escritas de manera vívida y vivida por la inmediatez de haberlo hecho en las aulas de la escuela. El primer día consignado es el 16 de noviembre de 1918 (lleva un mes dando clase en el Instituto-Escuela) y el último data del 4 de octubre de 1959, en el que refiere su profundo pesar por tener que jubilarse. El diario, que no es «diario» pues omite días, meses, incluso años, refleja muy claramente su tono vital y el grado de entusiasmo, ambos muy distintos en cada una de las tres etapas en que podría dividirse: aquella de los últimos años de la Monarquía, los tiempos de la II República hasta la terrible interrupción de la vida española durante la guerra civil y el periodo posterior a la guerra: «No ha llegado la paz, ha llegado la victoria».

El conjunto de reflexiones de tipo más teórico, titulado «Los problemas de la escuela» está formado por una colección de artículos publicados en el *Boletín de la Institución Libre de Enseñanza*, entre los años 1932 y 1936.

La biografía de María Sánchez Arbós, siempre marcada por una fuerza vital extraordinaria, es intensa en hechos y conceptos y extensa en su localización geográfica. Desde sus primeros estudios de maestra no dejó de formarse y estudió Filosofía y Letras; enlazó con la Institución Libre de Enseñanza, donde encontró el ambiente más propicio donde dar cauce a sus inquietudes; obtuvo un beca para ampliar estudios en la Residencia de Señoritas; opusculó brillantemente en varias ocasiones; colaboró con el Ministerio de Instrucción Pública, en Huesca, Zaragoza, Madrid, La Granja de San Ildefonso, La Laguna, Valencia, Daganzo.

Si algo caracteriza los diarios de esta maestra pionera (Huesca, 31 de octubre de 1889-Madrid, 15 de agosto de 1976) es la combinación de humildad, sinceridad, sencillez, cariño, duda metódica, entrega, entusiasmo, generosidad, capacidad de sacrificio, autoexigencia y exigencia y un gran amor y comprensión hacia el niño. Son estas las cualidades que las páginas de María Sánchez Arbós destilan y envuelven al lector y que definen la personalidad de su autora, quien siempre estuvo en activo excepto el tiempo que sufrió la depuración, la cárcel y la inhabilitación como funcionaria pública, periodo éste durante el cual tampoco dejó de llevar a cabo su labor docente bien como profesora particular, bien en colegios de titularidad privada, de los que consigna terrible desánimo que le causan. En 1952 fue rehabilitada para el magisterio y se jubiló en su último destino que fue el madrileño Instituto Isabel la Católica.

Son muchos los aspectos de este libro que pueden sorprendernos hoy. Cabe destacar, por un lado, su lenguaje sencillo y directo. Por otro, su modernidad en el hacer magisterial y sus opiniones sobre el *qué* y el *cómo* enseñar, coincidentes en su mayor parte con la concepción

actual del mundo escolar: planteamientos; programaciones flexibles; currículos adaptados a la condición socioeconómica; participación de las familias en la marcha de la escuela; organización de bibliotecas escolares; uso y abuso del libro de texto; coeducación...

*Mi diario* no se publicó hasta 1960. Salió a la luz en México (en España su publicación era impensable) y fueron impresos tan solo 100 ejemplares. Hoy, ve la luz por segunda vez en nuestro país (la primera edición del año 2000 no es tan completa): lectura imprescindible para todos aquellos que nos hemos dedicado, nos dedicamos o nos vamos a dedicar a la tarea de educar –«orientar», como dice la autora– en nuestro país, pues no es posible la comprensión del presente sin habernos asomado al balcón construido en el pasado por nuestros aquellos que nos antecedieron.

**Ana Franco**



# NORMAS PARA LA PRESENTACIÓN DE ORIGINALES

## Revista de Educación (Madrid)

ISSN: 0034-8082

ISSN (Internet): 1988-592X

### 1. Trayectoria, cobertura y contenido

La *Revista de Educación (Madrid)*, publicada por el Ministerio de Educación y Ciencia español, es una **revista científica arbitrada**, siendo objeto de evaluación externa todos los trabajos (recibidos y por encargo). De periodicidad cuatrimestral, fue fundada en 1940 (con el título de *Revista Nacional de Educación*), siendo la actual *Revista de Educación* continuación de aquella (con este nombre desde 1952). Acepta para su publicación principalmente trabajos originales de investigación básica y aplicada, así como experiencias de innovación sistematizadas, ensayos, informes, y recensiones de publicaciones recientes de relevancia en el campo de la educación. Su objetivo es difundir el conocimiento en esta área para la mejora de la investigación, la política y la práctica educativa, siendo sus destinatarios la comunidad científica en educación, administradores y profesores. Las secciones de la revista son: **a) Monográfico; b) Investigaciones y estudios; c) Informes y ensayos; d) Experiencias educativas (innovación); e) Recensiones**. Se aceptan trabajos originales en español y en inglés.

### 2. Presentación y envío de artículos

*Manual de estilo: Publication Manual of the American Psychological Association (APA)*, 5ª ed. Washington: APA, 2001 (consultar Normas APA en [www.revistaeducacion.mec.es](http://www.revistaeducacion.mec.es) o en [www.apastyle.org](http://www.apastyle.org)).

*Soprote y formato:* se remitirá una copia del original impresa (en DIN-A4), con el texto en procesador Word) y una copia de archivo en CD-Rom o disquete.

*Datos, contenido, estructura y estilo del artículo:* los autores observarán las siguientes normas para lograr una mayor eficacia en la gestión editorial de los trabajos:

- **Datos de autoría y, en su caso, de la financiación de la investigación:** en hoja aparte se indicarán necesariamente:
  - Datos del autor o autores: el nombre<sup>2</sup> y los dos apellidos de cada autor; con el/los grado/s académico/s más alto/s y la filiación institucional completa, esto es: nombre completo del centro y del departamento y/o institución subordinada a las que pertenezcan los autores, así como el nombre, la dirección postal, el teléfono y el correo electrónico del autor que se designe responsable de la correspondencia sobre el artículo.
  - Los autores están obligados a declarar, si este es el caso, el apoyo o financiación recibida para realizar la investigación que se pretende publicar; así como los proyectos de investigación o contratos financiados de la que es resultado. Se recomienda incluir al principio del artículo un pie de página donde se haga constar el agradecimiento por las ayudas recibidas.

<sup>(1)</sup> Los autores debe tener en cuenta que al publicar la Revista mayor proporción de investigaciones originales, el resto de los artículos (ensayos, informes y experiencias) pueden tener más demora en su publicación.

<sup>(2)</sup> **Aviso importante para los autores:** se recomienda la adopción de un *nombre de pluma*, para una correcta indización del artículo en las bases internacionales, que consista en un Nombre y un solo Apellido para nombres y apellidos poco comunes, o bien el Nombre y los dos Apellidos unidos por un guión para los más corrientes (Ej. María Clein-Acosta). Para más información véase [www.accesowok.fecyt.es](http://www.accesowok.fecyt.es) (enlace a «sabes qué tienes que hacer para identificar tus publicaciones científicas»).

• **Presentación del artículo:**

- Se omitirá toda referencia al nombre del autor o autores del trabajo y a sus credenciales.
- El artículo ha de estar encabezado por:
  - El **título del trabajo**, lo más ilustrativo y conciso posible, escrito primero en español y después en inglés, conteniendo entre 8-9 palabras clave significativas, extraídas, a ser posible del Tesoro de ERIC (Education Resources Information Center), que se encuentra accesible en Thesaurus – [www.eric.ed.gov](http://www.eric.ed.gov).
  - Un **resumen en español**, que debe tener **300 palabras** y, a continuación, la traducción de este al inglés (*abstract*). El resumen y el *abstract* serán publicados tanto en la edición impresa como electrónica de la revista. El resumen debe estructurarse de la siguiente manera (de acuerdo con el formato IMRYD): **introducción**, que recogerá el objetivo o finalidad de la investigación; **metodología**: incluirá los procedimientos básicos (diseño, selección de muestras o casos, métodos y técnicas de experimentación u observación y de análisis); **resultados**: principales hallazgos (dar datos específicos y su significación estadística, cuando corresponda); y **discusión o conclusiones**<sup>3</sup>.
  - Debajo del resumen, se deben incluir **de 5 a 10 palabras clave** o frases cortas (lexemas o descriptores), que también irán expresadas **en español y en inglés**. Se usarán palabras clave o términos internacionalmente aceptados en el campo de la educación para expresar conceptos y contenidos (véase Tesoro de ERIC).

• **Texto del artículo:**

- **Extensión**: en el caso de investigaciones o estudios la extensión no sobrepasará las 8.000 palabras. En el caso de experiencias educativas, ensayos e informes no sobrepasará las 4.000 palabras. En todos los casos, dicha extensión incluirá notas, referencias, bibliografía y elementos gráficos. La versión completa del artículo será publicada en la edición electrónica de la Revista.
- **Estructura**: en el caso de investigaciones y estudios, se recomienda que el artículo contemple, al menos, los siguientes aspectos: planteamiento del problema o tema objeto de estudio, antecedentes y fundamentación teórica, diseño y metodología, resultados, discusión de resultados, conclusiones, limitaciones del estudio y, en su caso, prospectiva.
- **Nombres, símbolos y nomenclatura**: los autores deben emplear aquellos que estén normalizados para cada disciplina.
- **Citas textuales**: las citas textuales deberán destacarse entre comillas y a continuación, entre paréntesis, el apellido e iniciales del autor; año de publicación y páginas de las que se ha extraído dicho texto.
- **Los esquemas, dibujos, gráficos, tablas, fotografías**, etc. deben ser aquellos necesarios para complementar o clarificar el texto. Se numerarán consecutivamente en función del tipo (tabla, gráfico...), se insertarán en el lugar idóneo dentro del cuerpo del texto del artículo y se presentarán en blanco y negro. Los gráficos, esquemas y tablas deberán presentarse en un formato que no sea imagen con el fin de facilitar las modificaciones posteriores si fuese necesario en la maquetación del artículo.
- **Las notas** se numerarán consecutivamente y su texto se recogerá a pie de página, restringiéndolas al mínimo necesario. Se evitarán las notas que sean simples referencias bibliográficas, en cuyo caso deberán ir en el texto señalando solo el autor/es y, entre paréntesis, el año de publicación; la referencia completa se incluirá en las referencias bibliográficas.
- Al final del trabajo se incluirá una lista denominada **«Referencias bibliográficas»**; la veracidad de las citas bibliográficas serán responsabilidad del autor o autores del artículo. Estas serán presentadas por orden alfabético y deberán ajustarse a las normas APA<sup>4</sup> (en el Anexo IV de estas Normas se ofrece un extracto). Las citas bibliográficas que se hagan en el texto del artículo deben tener su correspondencia en las «referencias bibliográficas».

• El autor o autores deben adjuntar una **Carta de presentación y cesión de derechos**, donde harán constar:

- El título del artículo completo.
- Nombre/s del autor/es.
- Dirección del responsable de la correspondencia.
- Solicitud de evaluación del artículo.

<sup>3)</sup> La importancia de redactar un título y resumen adecuados estriba en que de su lectura dependerá en buena medida que los potenciales lectores lleguen a leer el trabajo completo (especialmente en búsquedas electrónicas).

<sup>4)</sup> Consultar en [www.revistaeducacion.mec.es](http://www.revistaeducacion.mec.es) o en [www.apastyle.org](http://www.apastyle.org).

- Justificación de la selección de la *Revista de Educación*.
  - Declaración de autoría: en el caso de trabajos realizados por más de un autor, todos los autores deben certificar que los firmantes han contribuido directamente al contenido intelectual del trabajo, que se hacen responsables del mismo, lo aprueban y están de acuerdo en que su nombre figure como autor.
  - Declaración de que el manuscrito es inédito, y no se encuentra en proceso de evaluación en ninguna otra publicación.
  - Sección de la revista en la que desea publicar el artículo, acorde con su naturaleza: a) Investigaciones o estudios; b) Ensayos o informes; o c) Experiencias educativas (innovación).
  - Cesión de derechos (*copyright*) del artículo a la *Revista de Educación*.
  - Declaración de posibles conflictos de intereses: los autores deben hacer declaración de aquellas actividades –especialmente las relaciones financieras– que pudieran introducir sesgos en los resultados del trabajo.
- **Envío de artículos:** debe hacerse por correo postal a la siguiente dirección:  
*Revista de Educación*  
 Instituto de Evaluación (MEPSYD)  
 C/ San Fernando del Jarama, 14, 1ª planta.  
 (28002 Madrid)

### 3. Proceso editorial

- *Recepción de artículos.* Se acusará su recibo, tras lo que la Secretaría de la Revista efectuará una primera valoración editorial consistente en comprobar: a) la adecuación al ámbito temático y el interés del artículo en función de los criterios editoriales de la Revista, y b) cumplimiento de los requisitos de presentación formal exigidos en las normas de publicación. La recepción del artículo no supone su aceptación.
- *Sistema de revisión por pares (per review).* Comprobados el cumplimiento de los requisitos formales y la adecuación al interés temático de la Revista, el artículo será enviado a evaluación por parte de dos o más revisores expertos (externos o del Consejo Asesor), de forma confidencial y anónima (doble ciego), quienes emitirán un informe sobre la conveniencia o no de su publicación, que será tomado en consideración por la Secretaría del Consejo Editorial. El protocolo de evaluación utilizado por los revisores se hace público como anexo a estas Normas (Anexo I) en el sitio web de la Revista. En el caso de juicios dispares entre los dos evaluadores, el trabajo será remitido a un tercer evaluador: El trabajo revisado que se considere puede ser publicado condicionado a la inclusión de modificaciones, deberá ser corregido y devuelto por los autores a la Revista en el plazo máximo de un mes, tanto si se solicitan correcciones menores como mayores. De ser necesario, la nueva versión será enviada de nuevo a los revisores externos, procedimiento que se seguirá hasta su definitiva aceptación por la Revista. Los autores recibirán los informes de evaluación de los revisores, de forma anónima, para que éstos puedan realizar (en su caso) las correcciones oportunas.
- *Criterios de selección de revisores.* La selección de revisores es competencia de los editores de la revista, quienes tienen en cuenta sus méritos académicos, científicos y la experiencia profesional, incluyendo especialistas tanto de origen nacional como internacional. Entre los revisores podrán figurar ocasionalmente miembros del Comité Asesor. La Revista cuenta con revisores especialistas para contrastar los procedimientos metodológicos empleados en los trabajos.
- *Criterios de política editorial.* Los factores en los que se funda la decisión sobre la aceptación-rechazo de los trabajos por parte de los editores de la Revista son los siguientes: a) Originalidad: totalmente original, información valiosa, repetición de resultados conocidos; b) Actualidad y novedad; c) Relevancia: aplicabilidad de los resultados para la resolución de problemas concretos; d) Significación: para el avance del conocimiento científico; e) Fiabilidad y validez científica: calidad metodológica contrastada; f) Presentación: buena redacción, organización (coherencia lógica y presentación material).
- *Proceso de publicación.* Una vez finalizado el proceso de evaluación, se enviará al autor principal del trabajo la notificación de aceptación o rechazo para su publicación. Asimismo será informado de la llegada de las primeras pruebas de imprenta para facilitar su revisión dentro del plazo establecido por la Revista.
- *Permiso para reproducir material publicado.* El contenido de los trabajos puede ser reproducido, total o parcialmente, citando procedencia y solicitando autorización escrita al Editor antes de la publicación de dicho material.

- *Principios éticos relativos a la investigación y la publicación.* Es obligación de la *Revista de Educación* detectar y denunciar las siguientes prácticas deshonestas sobre los diversos supuestos de fraude científico: a) fabricación, falsificación u omisión de datos y plagio; b) Publicación duplicada. c) Autoría y conflictos de interés (véase apartado 4 de estas normas).
- Al citar trabajos publicados en esta Revista se debe incluir siempre la siguiente información: *Revista de Educación (Madrid)*, número de la revista, página/s y año de publicación.

#### **4. Responsabilidades éticas**

- La revista no acepta material previamente publicado. Los autores son responsables de obtener los oportunos permisos para reproducir parcialmente material (texto, tablas o figuras) de otras publicaciones y de citar su procedencia correctamente. Estos permisos deben solicitarse tanto al autor como a la editorial que ha publicado dicho material.
- En la lista de autores firmantes deben figurar únicamente aquellas personas que han contribuido intelectualmente al desarrollo del trabajo.
- La Revista espera que los autores declaren cualquier asociación comercial que pueda suponer un conflicto de intereses en conexión con el artículo remitido.
- Los autores deben mencionar en la sección de métodos que los procedimientos utilizados en los muestreos y controles han sido realizados tras obtención de un consentimiento informado.
- Se adjuntará al artículo una *Hoja de Identificación y Cesión de Derechos* (facilitada por la revista) firmada por todos los autores. La *Revista de Educación* se reserva la facultad de introducir las modificaciones que considere oportunas en la aplicación de estas normas, de lo que se informará oportunamente. Los originales enviados no serán devueltos.

**Anexo I.** - PROTOCOLO DE EVALUACIÓN A UTILIZAR POR LOS REVISORES (consultar en *Normas para la presentación de originales*: [www.revistaeducacion.mec.es](http://www.revistaeducacion.mec.es)).

**Anexo II.** - LISTA DE COMPROBACIÓN PARA LOS AUTORES (consultar en *Normas para la presentación de originales*: [www.revistaeducacion.mec.es](http://www.revistaeducacion.mec.es)).

**Anexo III.** - CARTA DE PRESENTACIÓN Y CESIÓN DE DERECHOS (consultar en *Normas para la presentación de originales*: [www.revistaeducacion.mec.es](http://www.revistaeducacion.mec.es)).

**Anexo IV.** - EJEMPLIFICACIÓN REDUCIDA DE LAS NORMAS APA (5ª edición) (consultar en *Normas para la presentación de originales*: [www.revistaeducacion.mec.es](http://www.revistaeducacion.mec.es)).

**Se agradece de antemano a los autores la observación de estas normas, lo que permitirá agilizar el proceso de evaluación y edición.**

## Normas para la presentación de Recensiones

1. Con relación a la obra reseñada:
  - Preferentemente publicaciones científicas o innovaciones relevantes en el campo de la educación.
  - Relevancia del tema objeto de la obra.
  - Especialización y prestigio del autor o autores, y de la editorial.
  - Actualidad (año de publicación).
  - Dado el carácter científico de la revista, se excluyen reseñas de obras de divulgación, libros de texto y materiales exclusivamente didácticos, o de cualquier obra que defienda o propugne principios contrarios a los valores democráticos.
  
2. Con relación a la reseña:
  - Calidad del texto y especialización del recensor en el tema de la obra.
  - La recensión no debe sobrepasar las 700 palabras.
  - La reseña estará encabezada por los datos de la obra reseñada (siguiendo estrictamente las normas de la Revista).
  - El nombre del autor de la recensión deberá escribirse al final del texto.

La recensión se enviará a la dirección siguiente: [redaccion.revista@mepsyd.es](mailto:redaccion.revista@mepsyd.es). El autor o autora enviará también, en hoja aparte, los siguientes **datos personales**: nombre y apellidos, dirección postal, teléfono y dirección de correo electrónico.

**Le rogamos que, en caso de que la *Revista de Educación (Madrid)* le haya facilitado la obra a reseñar, nos sea devuelta, pues pertenece a los fondos de la biblioteca del Ministerio de Educación y Ciencia.**



# General Guidelines for the Submission of Manuscripts

## Revista de Educación (Madrid)

ISSN: 0034-8082  
ISSN (Internet): 1988-592X

### 1. Experience, Coverage and Content

*Revista de Educación (Madrid)*, edited by the Spanish Ministry of Education and Science, is an **arbitrated scientific publication** whose papers, both received and commissioned to, are all subject to external assessment. Four monthly published, it was founded in 1940 with the title *Revista Nacional de Educación*. The current *Revista de Educación (Madrid)*, in continuation with that one, is published with this title since 1952. It publishes basic and applied research, as well as systematized innovation experiences, essays and reports and reviews of recent publications of educational interest. Its main objective is the dissemination of knowledge in this area for the improvement of research and educational policy and practice. Its target is the scientific community in education, administrators and teachers. The main sections of the Journal are the following<sup>1</sup>: **a) Monographic; b) Research Projects and Studies; c) Reports and Essays; d) Educational Experiences (innovation); e) Reviews**. Original papers in Spanish and English are accepted.

### 2. Submissions of Manuscripts and Style of Presentation

*Style Manual: Publication Manual of the American Psychological Association (APA)*. 5<sup>th</sup> ed. Washington: APA, 2001 (more information about APA criteria in [www.revistaeducacion.mec.es](http://www.revistaeducacion.mec.es) or [www.apastyle.org](http://www.apastyle.org)).

*Medium and Format*: Provide 1 copy of the original manuscript (on DIN-A4 paper and using Word as a text processor) and a file copy stored on a CD-Rom o diskette.

*Data, content, structure and style of the paper*; authors will take into account the following guidelines to achieve a greater efficacy in the editorial management of manuscripts:

- **Authorship data and, when appropriate, research funding data**. It should be specified on a separate sheet of paper:
  - Personal and professional details: author or authors forename<sup>2</sup> and first surname, highest university degree/s and full institutional affiliation, i.e. name of the Faculty and the Department and/or subordinate institution to which the authors belong to, as well as full name, telephone number, full postal address and email address of the author responsible for correspondence.
  - If it is the case, authors have to declare any grant or financial support received for the carrying out of the research intended to be published, as well as those research projects or funding contracts from which it results from. Authors are requested to acknowledge the grants or financial support received in a footer included at the beginning of the paper.

<sup>(1)</sup> Authors should consider that the Journal publishes more original research projects, so the rest of manuscripts (essays, reports and experiences) could be published later.

<sup>(2)</sup> **Important notice to authors**: It is advisable to adopt a **writing name** for the correct indexing of papers according to international bases. It should consist of the first and last name for not very common names and surnames, or the name and two surnames connected by a hyphen for not very common names (Ex. María Clein-Acosta). For further information see [www.accesowok.fecyt.es](http://www.accesowok.fecyt.es) (link «*Sabes qué tienes que hacer para identificar tus publicaciones científicas?*») (Do you know what you have to do to identify your scientific papers?).

- **Style of presentation:**
  - Any reference to the author(s) will be omitted from the paper.
  - The paper should be headed by:
    - The **title**. It should be concise but illustrative. Written first in Spanish and then in English. Include 8-9 relevant key words and, if it is possible, taken from the ERIC Thesaurus (Education Resources Information Centre), which is available in the following email address: [www.eric.es.gov/thesaurus](http://www.eric.es.gov/thesaurus).
    - An **abstract in Spanish**, which must contain **300 words**, followed by its translation into English. Both, Spanish and English abstracts will be published in the printed and electronic versions of the journal. The summary should be structured in the following way (IMRYD format): **Introduction**, which should include the main objective of the research; **methodology**, which should include basic procedures (design, selection of samples or cases, methods and observation/ experimental and analysis techniques); **results**: Main finds (essential data and their statistical relevance should be provided, as appropriate); and **main conclusions and discussion**<sup>3</sup>.
    - **5-10 key words** or short sentences (lexemes or descriptors) should be provided immediately after the summary, which will also be written both in **Spanish and English**. Key words or internationally accepted words in the educational field to express concepts and contents (see ERIC Thesaurus) will be used.
  
- **Text:**
  - **Length:** Research projects or studies should not exceed 8,000 words. Educational experiences, essays and reports should not exceed 4,000 words. In all cases, notes, references, bibliography and graphic elements will be included in the aforementioned length. The full version of the paper will be published in the electronic edition of *Revista de Educación*.
  - **Structure:** In the case of research projects and studies it is recommended that the paper should include, at least, the following aspects: Description of the problem or subject matter under analysis, prior work and theoretical bases, organization and methodology, results, limitations, conclusions and further developments, if appropriate.
  - **Names, symbols and nomenclature:** Authors should use those normalized for every discipline.
  - **Direct citations:** They should be distinguished from the rest by means of inverted commas, and followed by the author's last name and first name initials, year of publication and pages from where the text has been extracted; all of which should appear in brackets.
  - **Diagrams, pictures, figures, tables, photographs**, etc. should be those necessary to complete or clarify the text. They should be numbered consecutively depending on the type (table, graphic...); they will be inserted in a suitable place within the body of the text of the article and must be submitted in black and white. Graphics, diagrams and tables should not be images, so that further changes, if needed, may be possible.
  - **Notes** should be numbered consecutively and the information corresponding to them should be printed at the bottom of each page, restricted until maximum. Simple bibliography references have to be avoided; in that case they will be inserted in the text mentioning the author and the year of publication in brackets. Then, the complete reference will be included in the bibliography.
  - At the end of the paper, a list called 'Bibliographical References' will be included. The veracity of bibliographical citations will be the author or authors' responsibility. They will be presented alphabetically and they should adjust to APA criteria<sup>4</sup> (a summary of these criteria is provided in Annex IV). References in the article should have their correspondence in the 'Bibliographical References'.
  
- The author or authors should submit a **Presentation and transmission of author's rights letter** where they should include the following data:
  - Full title of the paper.
  - Author(s)' name(s).
  - Full postal address of the author responsible for correspondence.
  - Request for the review of the paper.

<sup>3</sup> The importance of writing an appropriate title and summary lies in the fact that its reading will depend, to a great extent, on the possibility that potential readers decide to read the full paper (especially as regards electronic searches).

<sup>4</sup> Further information in [www.revistaeducacion.mec.es](http://www.revistaeducacion.mec.es) or in [www.apastyle.org](http://www.apastyle.org).

- Justification of the selection by the *Revista de Educación*.
- Declaration of authorship: In the case of those papers carried out by more than one author, all of them should certify that the signatories to it have all directly contributed to the intellectual content of the paper; that they are responsible for it, they approve of it and agree for their name to appear as one of the main authors.
- Declaration that the paper is unpublished and it is not being assessed by any other publication.
- Section of the journal where the paper should be published, according to its nature: a) Research projects or Studies; b) Essays or Reports; or c) Educational experiences (innovation).
- Transmission of author's rights (copyright) to *Revista de Educación*.
- Declaration of possible conflicting interest: The Journal expects authors to declare any financial relationship that might suppose a conflict of interests connected to the results of the article submitted.

- **Sending of papers:**

They should be sent by normal post to the following address:

Revista de Educación (Madrid)  
 Instituto de Evaluación (MEPSYD)  
 C/ San Fernando del Jarama, 14, 1ª planta.  
 (28002 Madrid)  
 SPAIN

### 3. Editorial Process

- *Reception of papers.* Receipt of papers will be acknowledged. Then, the Secretarial Department of *Revista de Educación* will carry out a first editorial assessment consisting of: Checking a) their adequacy to the corresponding thematic field and the interest of the paper according to the editorial criteria of the Journal; and b) fulfilling of style presentation requirements stated in publication guidelines. The reception of the paper implies its acceptance.
- *Review System peer review.* Once the fulfilling of formal requirements and the adequacy to the corresponding thematic field of the Journal have been checked, the paper is confidentially and anonymously reviewed by two experts (external or belonging to the Editorial Advisory Board); this is done according to the double-mask method. These experts will provide a report on the convenience or not for the publication of the paper, which will be taken into consideration by the Secretarial Department of the Journal. The assessment protocol used by experts is made public as an annex to these Guidelines (Annex I) in the Journal website. In case the two reviewers provide different assessment, the paper will be sent to a third expert. Then, those papers which are considered as worth publishing, conditional on the inclusion of modifications, will be corrected and returned to the Journal in one month, both if major or minor corrections are requested. If it is the case, the new version will be sent again to external correctors, and this procedure will be followed until its definite acceptance by the Journal. Authors will be sent the assessment report issued by experts in an anonymous way, so that they can make in any case the corresponding corrections.
- *Selection criteria of experts.* The editors of the Journal are responsible for the selection of experts, and they take into account their academic and scientific merits, and their professional experience, including both national and international specialists. As an exceptional case, members of the Journal's Advisory Council will occasionally act as experts. The Journal has experts to verify the methodological procedures used in papers.
- *Editorial policy criteria.* The factors which influence the decision on the acceptance-rejection of papers by the editors of the Journal are the following: a) Originality: Totally original, valuable information, repetition of known results; b) Topicality and innovation; c) Relevance: Applicability of results for the solution of specific problems; Significance: For the advance of scientific knowledge; e) Scientific reliability and validity: Contrasted methodological quality; f) Presentation: Good writing, organization (logical coherence and material presentation).
- *Publication process.* Once the assessment process has come to an end, the paper's main author will receive a formal notification stating the acceptance or rejection of its publication. Likewise, s/he will be informed about the carrying out of the first proofs/ galley proofs so as to facilitate its revision within the deadline established by the Journal.

- *Permission to reprint the published material.* The content of papers may be total or partially reprinted. For this purpose, the origin of the paper will be specified and the editor will be provided a written permission by the author before the paper is published.
- *Ethical principles as regards research and publication.* It is the *Revista de Educación* obligation to detect and report the following dishonest practices concerning the several assumptions on scientific fraud: a) Fabrication, forging or omission of data and plagiarism; b) duplicated publication; c) authorship and conflicting interest (see section 4 of these guidelines).
- When citing the papers published in this Journal, the following information should be included: *Revista de Educación*, number of the journal, paper/s and year of publication.

#### **4. Ethical Responsibilities**

- The journal does not accept previously published material. Authors are responsible for obtaining the appropriate permission for partial reproductions of material (text, tables or figures) from other publications. This permission must be sought from both author and publisher of the material.
- Only those persons who have contributed intellectually to the development of the study can appear on the list of authors.
- The journal expects authors to declare any personal or financial relationship that might suppose a conflict of interests connected to the article submitted.
- Authors should indicate in the section devoted to methods that the procedures used in the sampling and controls have been carried out after the obtaining of an informed permission.
- Authors will enclosed an *Identification and Transmission of Rights Sheet* (provided by the Journal) signed by all authors. *Revista de Educación* reserves the right to make any change or introduce modifications in the application of these norms, which will be reported at the proper time. Originals will not be returned.

**Annex I.** - ASSESSMENT PROTOCOL USED BY EXPERTS (consult in General Guidelines for the Submission of Manuscripts: [www.revistaeducacion.mec.es](http://www.revistaeducacion.mec.es)).

**Annex II.** - CHECK LIST FOR AUTHORS (consult in General Guidelines for the Submission of Manuscripts: [www.revistaeducacion.mec.es](http://www.revistaeducacion.mec.es)).

**Annex III.**- PRESENTATION AND TRANSMISSION OF AUTHOR'S RIGHTS LETTER (consult in General Guidelines for the Submission of Manuscripts: [www.revistaeducacion.mec.es](http://www.revistaeducacion.mec.es)).

**Annex IV** - BRIEF EXEMPLIFICATION OF APA GUIDELINES (5<sup>th</sup> edition) (consult in General Guidelines for the Submission of Manuscripts: [www.revistaeducacion.mec.es](http://www.revistaeducacion.mec.es)).

**Revista de Educación (Madrid) thanks the authors in advance for complying with these guidelines, which will speed the review and publishing process.**

## Norms and Criteria for Submitting Reviews

1. With regard to the work reviewed:
  - Preferably scientific publications or outstanding innovations in the field of education.
  - Relevance of the work's subject.
  - Specialization and prestige of the author or authors and of the publishing house.
  - Topicality (year in which it was published).
  - Due to the scientific nature of the journal, all the reviews related to works spreading information, textbooks and didactic materials are excluded, together with those connected with any work that defends or supports principles which go against democratic values.
  
2. With regard to the review:
  - Quality of the text and specialization of the reviewer in the subject.
  - It cannot exceed 700 words.
  - It must be headed by the details of the reviewed work (strictly following the norms of the journal).
  - The reviewer's name must be written at the end of the text.

The review must be sent to the following electronic address [redaccion.revista@mepsyd.es](mailto:redaccion.revista@mepsyd.es). The author must also send on a different piece of paper the following **personal details**: Postal address, telephone number and e-mail address.

**Please, in case you have been sent the work to be reviewed, you should send it back to the *Revista de Educación (Madrid)*, since it belongs to the list of titles of the Ministry of Education and Science.**

# SUMARIO DE LA SECCIÓN MONOGRÁFICA

## El valor añadido en educación

### Value-added in education

ROSARIO MARTÍNEZ ARIAS, JOSÉ LUIS GAVIRIA SOTO Y MARÍA CASTRO MORERA. **Concepto y evolución de los modelos de valor añadido en educación.**

ANDREW RAY, HELEN EVANS Y TANYA MCCORMARK. **El uso de los modelos nacionales de valor añadido para la mejora de las escuelas británicas.**

YEOW MENG THUM. **No Child Left Behind: retos metodológicos y recomendaciones para la medida del progreso anual adecuado.**

EVA BAKER. **Consideraciones de validez prioritaria para la evaluación formativa y de rendición de cuentas.**

MARÍA CASTRO MORERA, COVADONGA RUIZ Y ESTHER LÓPEZ. **Forma básica del crecimiento en los modelos de valor añadido: vías para la supresión del efecto de regresión.**

MARÍA EUGENIA FERRÃO. **Sensibilidad de las especificaciones de los modelos de valor añadido: midiendo el estatus socioeconómico.**

JOSÉ LUIS GAVIRIA SOTO, M<sup>a</sup> CHANTAL BIENCINTO, ENRIQUE NAVARRO. **Invarianza de la estructura de covarianzas de las medidas de rendimiento académicos en estudios logitudinales en la transición de la Educación Primaria a Secundaria.**

LUIS LIZASOAIN HERNÁNDEZ Y LUIS JOARISTI OLARIAGA. **Análisis de la dimensionalidad en los modelos de valor añadido: estudio de las pruebas de matemáticas empleando técnicas factoriales y métodos no paramétricos basados en TRI (Teoría del Respuesta al Ítem).**

ÁNGELES BLANCO, CORAL GONZÁLEZ Y XAVIER ORDOÑEZ. **Efectos de la complejidad del rasgo latente en la estimación del rendimiento académico. Un estudio longitudinal.**

ROSARIO MARTÍNEZ ARIAS. **Usos, aplicaciones y problemas de los modelos de valor añadido en educación.**



Edición completa en  
[www.revistaeducacion.mepsyd.es](http://www.revistaeducacion.mepsyd.es)