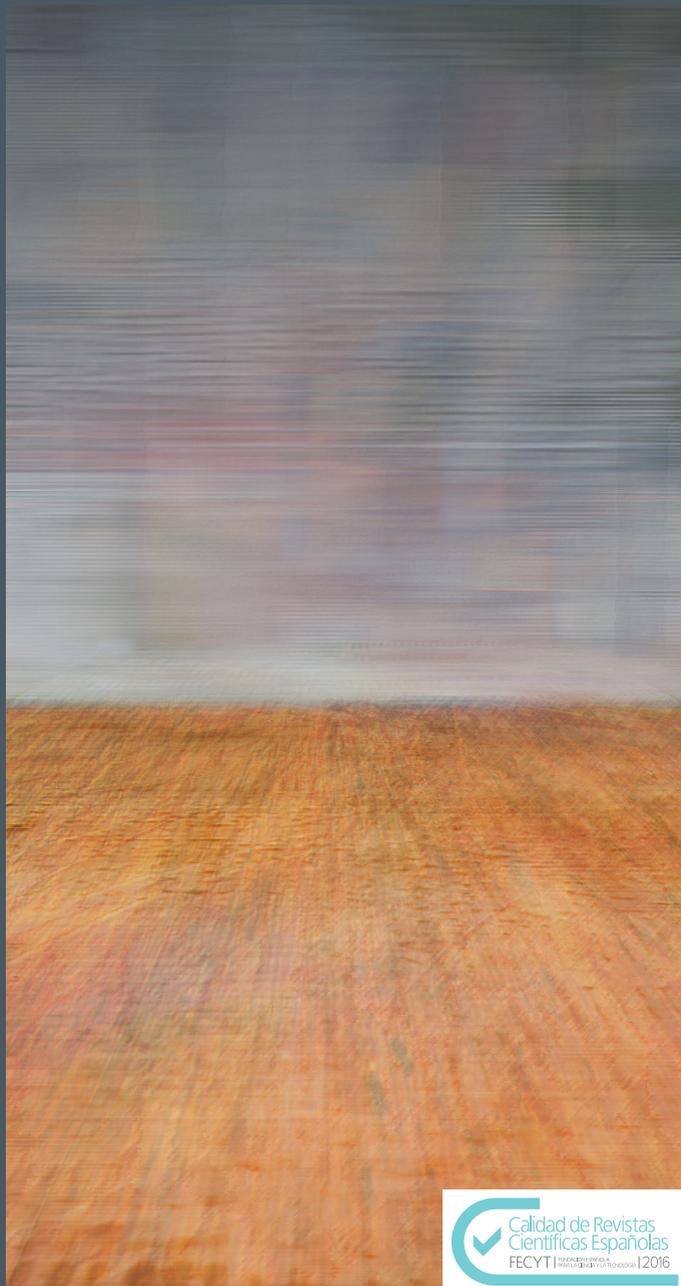


# revista de **e**DUCCIÓN

Nº 381 JULIO-SEPTIEMBRE 2018



GOBIERNO  
DE ESPAÑA

MINISTERIO  
DE EDUCACIÓN, CULTURA  
Y DEPORTE



Calidad de Revistas  
Científicas Españolas  
FECYT | INSTITUTO ESPAÑOL DE INVESTIGACIONES CIENTÍFICAS | 2016

**revista de**  
**EDUCACIÓN**



**N° 381 JULIO-SEPTIEMBRE 2018**

# revista de EDUCACIÓN

**Nº 381 Julio-Septiembre 2018**

Revista trimestral

Fecha de inicio: 1952



## **MINISTERIO DE EDUCACIÓN, CULTURA Y DEPORTE**

SECRETARÍA DE ESTADO DE EDUCACIÓN, FORMACIÓN PROFESIONAL Y UNIVERSIDADES

Instituto Nacional de Evaluación Educativa  
Ministerio de Educación, Cultura y Deporte  
Paseo del Prado, 28, 4.ª planta  
28014 Madrid  
España

Edita

© SECRETARÍA GENERAL TÉCNICA

Subdirección General de Documentación y Publicaciones.

Catálogo de publicaciones del Ministerio: [mecd.gob.es](http://mecd.gob.es)

Catálogo general de publicaciones oficiales: [publicacionesoficiales.boe.es](http://publicacionesoficiales.boe.es)

Edición: 2018

NIPO línea: 030-15-016-X

NIPO ibd: :030-15-017-5

ISSN línea: 1988-592X 0034-8082

ISSN papel: 0034-8082

Depósito Legal: M.57/1958

Diseño de la portada: Dinarte S.L.

Maqueta: Solana e hijos, Artes Gráficas S.A.U.

## CONSEJO DE DIRECCIÓN

### PRESIDENTE

Marcial Marín Hellín

Secretario de Estado de Educación, Formación Profesional y Universidades

### VOCALES

Jorge Sainz González

Secretario General de Universidades

Marco A. Rando Rando

Director General de Evaluación y Cooperación Territorial

Rosalía Serrano Velasco

Directora General de Formación Profesional

\_\_\_\_\_

Director General de Política Universitaria

Mónica Fernández Muñoz

Secretaría General Técnica

Carmen Tovar Sánchez

Directora del Instituto Nacional de Evaluación Educativa

Amparo Barbolla Granda

Subdirectora General de Documentación y Publicaciones

Violeta Miguel Pérez

Directora del Centro Nacional de Innovación e Investigación Educativa

## CONSEJO EDITORIAL

### DIRECTOR

Carmen Tovar Sánchez

### EDITOR JEFE

José Luis Gaviria Soto

### EDITOR ADJUNTO

David Reyero García

### VOCALES

Antonio Cabrales Goitia (U. Carlos III de Madrid); Caterina Casalmiglia (U. Autónoma de Barcelona); José Luis García Garrido (U. Nacional de Educación a Distancia); Antonio Lafuente García (CSIC); Leoncio López Ocón (CSIC); Clara Eugenia Núñez Romero (U. Nacional de Educación a Distancia); Arturo de la Orden Hoz (U. Complutense de Madrid); Lucrecia Santibáñez (RAND Corporation); Javier Tourón Figueroa (U. de Navarra); Pablo Zoido (OCDE)

## REDACCIÓN

**Colaboradores externos:** Jorge Mañana Rodríguez y Jesús García Laborda

## ASESORES CIENTÍFICOS

### Internacional

Aaron Benavot (State University of New York SUNY-Albany); Abdeljalil Akkari (Profesor de la Universidad de Ginebra, Suiza); Jorge Baxter (Organización de Estados Americanos); Mark Bray (University of Hong Kong); 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 (Profesor del Departamento de Educación, Universidad Autónoma de Aguascalientes, 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); Sagrario Avezuela Sánchez (IES Lázaro Cárdenas, CAM); 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 Dosal (U. de Santiago); Gerardo Echeita (U. Autónoma de Madrid); Juan Manuel Escudero (U. de Murcia); Mariano Fernández Enguita (U. Complutense de Madrid); Joaquín Gairín (U. Autónoma de Barcelona); M.ª Ángeles Galino; J. L. García Garrido (UNED); Daniel Gil (U. de Valencia); José Gimeno Sacristán (U. de Valencia); Fuensanta Hernández Pina (U. de Murcia); Carmen Labrador (U. Complutense); Ramón L. Facal (IES Pontepedriña, Santiago de Compostela. Revista Iber); Miguel López Melero (U. de Málaga); Carmen Maestro Martín (IES Gran Capitán, CAM); Carlos Marcelo (U. de Sevilla); 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); Juan Manuel Moreno Olmedilla (Banco Mundial); 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.ª Dolores de Prada (Inspección); Joaquim Prats (U. de Barcelona); Manuel de Puelles (UNED); Tomás Recio (U. de Cantabria); Luis Rico (U. de Granada); Enrique Roca Cobo; Juana M.ª Sancho (U. de Barcelona); Juan Carlos Tedesco (UNESCO); Alejandro Tiana Ferrer (UNED); Consuelo Uceda (Colegio La Navata, Madrid); Consuelo Vélaz de Medrano (UNED); 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, Cultura y Deporte español. Fundada en 1940, y manteniendo 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 en este campo, tanto desde una perspectiva nacional como internacional. La revista es editada por la Subdirección General de Documentación y Publicaciones, y actualmente está adscrita al Instituto Nacional de Evaluación Educativa de la Dirección General de Evaluación y Cooperación Territorial.

Cada año se publican cuatro números con tres secciones: Investigaciones, Ensayos y Reseñas. Uno de los números anuales podrá contar con una sección Monográfica con convocatoria pública en esta web. Todos los artículos enviados a las diferentes secciones están sometidos a evaluación externa. En el primer número del año se incluye, además, un índice bibliográfico, y en el segundo un editorial con la Memoria anual que recoge las principales estadísticas del proceso editor de ese período, la evolución de los índices de calidad e impacto, así como el listado de revisores externos.

Desde sus orígenes hasta 2006 la *Revista de Educación* se publicó en formato impreso. Desde 2006 se ha venido publicando en doble formato, impreso y electrónico. La edición impresa incluía 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 en la Redacción. La edición electrónica incluía los artículos y reseñas completos y es accesible a través de la página web ([www.mecd.gob.es/revista-de-educacion/](http://www.mecd.gob.es/revista-de-educacion/)), en la que además los lectores tienen acceso a otras informaciones de interés sobre la revista. Desde el segundo número de 2012 (358 mayo-agosto), la *Revista de Educación* se convierte en una publicación exclusivamente electrónica.

La *Revista de Educación* tiene un perfil temático generalista, pero solo evalúa, selecciona y publica trabajos enmarcados en un conjunto de líneas de investigación consolidadas, principalmente sobre: metodologías de investigación y evaluación en educación; políticas públicas en educación y formación; evolución e historia de los sistemas educativos; reformas e innovaciones educativas; calidad y equidad en educación; atención a la diversidad; currículo; didáctica; organización y dirección escolar; orientación educativa y tutoría; desarrollo profesional docente; cooperación internacional para el desarrollo de

la educación. Estas son las líneas de demarcación del perfil temático de la revista desde los años 60.

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

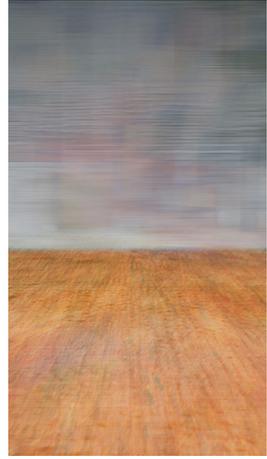
- *Bases de datos nacionales*: ISOC, BEG (GENCAT), PSICODOC, DIALNET, y REDINED (Red de Bases de Datos de Información Educativa).
- *Bases de datos internacionales*: Social Sciences Citation Index® (SSCI), Social Scisearch®, SCOPUS, Sociological Abstracts (CSA Illumina), PIO (Periodical Index Online, Reino Unido), IRESIE (México), ICIST (Canadá), hedbib (International Association of Universities - UNESCO International Bibliographic Database on Higher Education), SWETSNET (Holanda).
- *Sistemas de evaluación de revistas*: Journal Citation Reports/Social Sciences Edition (JCR), European Reference Index for the Humanities (ERIH), Latindex (Iberoamericana), scimago Journal & Country Rank (SJR), RESH, Difusión y Calidad Editorial de las Revistas Españolas de Humanidades y Ciencias Sociales y Jurídicas (DICE), carhus plus+, Matriu d'Informació per a l'Avaluació de Revistes (MIAR), Clasificación Integrada de Revistas Científicas (CIRC).
- *Directorios*: Ulrich's Periodicals Directory.
- *Catálogos nacionales*: Consejo Superior de Investigaciones Científicas (CSIC-ISOC), Red de Bibliotecas Universitarias (REBIUN), Centro Nacional de Innovación e Investigación Educativa (Ministerio de Educación, Cultura y Deporte), Catálogo Colectivo de Publicaciones Periódicas en Bibliotecas Españolas (Ministerio de Educación, Cultura y Deporte).
- *Catálogos internacionales*: WorldCat (USA), Online Computer Library Center (USA), Library of Congress (LC), The British Library Current Serials Received, King's College London, Catalogue Collectif de France (CCFr), Centro de Recursos Documentales e Informáticos de la Organización de Estados Iberoamericanos (OEI), COPAC National, Academic and Specialist Library Catalogue (Reino Unido), SUDOC Catalogue du Système Universitaire de Documentation (Francia), ZDB Zeitschriftendatenbank (Alemania).

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

## Investigaciones

CLAUDIA-ELENA VELÁZQUEZ-OLMEDO, LETICIA GALLEGOS-CÁZARES Y ELENA CALDERÓN-CANALES: Formación de representaciones intuitivas acerca del sonido en niños de preescolar .....	11
CAROLINE SHACKLETON: Vinculación de la prueba de comprensión oral del examen <i>CertAcles</i> de la Universidad de Granada con el MCER.....	37
FRANCISO JAVIER MURILLO y CYNTHIA MARTÍNEZ-GARRIDO: Incidencia de la crisis económica en la segregación escolar en España .....	67
MARÍA LUZ DIAGO EGAÑA, MARÍA JOSÉ CUETOS REVUELTA y PATRICIA GONZÁLEZ GONZÁLEZ: Análisis de las herramientas de medición de los Estilos de Aprendizaje .....	95
JOSETXU ORRANTIA, SARA SAN ROMUALDO, ROSARIO SÁNCHEZ, LAURA MATILLA, DAVID MUÑEZ y LIEVEN VERSCHAFFEL: Procesamiento de magnitudes numéricas y ejecución matemática.....	133
PATRICIA SUÁREZ y TRINIDAD SUÁREZ: Analfabetismo en el primer tercio del siglo XX: la creación de escuelas en Valdés (Asturias).....	155
XIMENA DUEÑAS HERRERA, ANA BOLENA ESCOBAR ESCOBAR, SILVANA GODOY MATEUS y JORGE LEONARDO DUARTE RODRÍGUEZ: Coaching a docentes y rendimiento académico: PTA en Colombia.....	181
ÁNGELES BLANCO-BLANCO: Estado de las prácticas científicas e investigación educativa. Posibles retos para la próxima década .....	207
SUSANA LÓPEZ SOLÉ, MIREIA CIVÍS ZARAGOZA y JORDI DÍAZ-GIBSON: Capital Social y Redes Sociales de Maestros: Revisión Sistemática .....	233

DAVID AGUILERA MORALES, TOBÍAS MARTÍN PÁEZ, VÍCTOR VALDIVIA RODRÍGUEZ, ÁNGELA RUIZ DELGADO, LETICIA WILLIAMS PINTO, JOSÉ MIGUEL VÍLCHEZ GONZÁLEZ y FRANCISCO JAVIER PERALES PALACIOS: La enseñanza de las ciencias basada en indagación. Una revisión sistemática de la producción española.....	259
RAQUEL FERNÁNDEZ-CÉZAR, NATALIA PINTO SOLANO y MARTA MUÑOZ HERNÁNDEZ: ¿Mejoran los proyectos de divulgación con experimentación la actitud hacia las clases de ciencias? .....	285
<b>Reseñas</b> .....	309



## **Investigaciones**



# Formación de representaciones intuitivas acerca del sonido en niños de preescolar<sup>1</sup>

## Construction of intuitive representations about sound in preschool children

DOI: 10.4438/1988-592X-RE-2017-381-379

Claudia-Elena Velázquez-Olmedo

Leticia Gallegos-Cázares

Elena Calderón-Canales

*Universidad Nacional Autónoma de México*

### Resumen

La presente investigación tuvo como objetivo conocer las representaciones intuitivas que tienen los niños de nivel preescolar acerca del sonido. El estudio se fundamenta en la teoría representacional, la cual estudia la construcción de conocimiento científico mediante el desarrollo de representaciones mentales, entendiendo por representación toda construcción, mediada cognitivamente por el sujeto, que le permite dar cuenta de un fenómeno o situación, y la teoría de la cognición corporal (embodiment), la cual hace referencia al papel del cuerpo en la construcción y explicitación de estas representaciones. Se realizó un estudio de casos constituido por seis niños de 3er grado de preescolar, los datos se obtuvieron mediante la aplicación individual de una entrevista semiestructurada basada en los temas: producción, percepción y propagación del sonido, que implicó el uso de material gráfico e instrumental lo que permitió contextualizar y delimitar las entrevistas, registradas en audio y video. Se realizaron paralelamente dos análisis diferentes, en el primero se construyeron categorías de análisis a partir de las respuestas verbales de los niños, en el segundo se identificaron y categorizan los gestos utilizados por los niños para explicar sus respuestas o en algunos casos siendo el gesto la propia respuesta. Los resultados revelan que los

---

<sup>(1)</sup> Esta investigación se llevó a cabo con el apoyo de la Secretaría de Educación Pública y el Consejo Nacional de Ciencia y Tecnología -México-(Ref. n° 240419. SEP-CONACyT). Agradecimientos a la Universidad Nacional Autónoma de México (UNAM) y al Instituto de Ciencias Aplicadas y Tecnología (ICAT).

niños tienen una representación sustancialista del sonido, es decir, consideran que el sonido tiene propiedades similares a las de los objetos (masa y trayectoria definida), basada en una intuición materialista acerca del sonido, que responde a experiencias fenomenológicas guiadas esencialmente por la percepción. Finalmente intuyen que existe una relación entre las vibraciones y el sonido pero no son capaces de hacer explícito de qué tipo de relación se trata.

*Palabras clave:* Educación de la primera infancia, Enseñanza de las ciencias, Cognición, Representación mental.

### **Abstract**

The present research aimed to know the intuitive representations that preschool children built about sound. The study is based on representational theory, which studies the construction of scientific knowledge through the development of mental representations, understanding by representation all construction, cognitively mediated by the subject, which allows him to account for a phenomenon or situation, and the theory of corporal cognition (embodiment), which refers to the role of the body in the construction and explanation of these representations. A case study was carried out consisting of six children of 3rd grade of preschool. The data were obtained through the individual application of a semi-structured interview based on the themes: production, perception and propagation of sound, which involved the use of graphic and material which allowed contextualizing and delimiting the interview. All the interviews were recorded in audio and video. Two different analyses were carried out in parallel, in the first, analysis categories were constructed from the verbal responses of the children. In the second analysis, the gestures used by the children to explain their answers were identified and categorized, or in some cases the gesture itself was the response. The results reveal that children have a substantialist representation of sound, that is, they consider that sound has similar properties to those objects (mass and defined trajectory), based on a materialistic intuition about sound, which responds to experiences phenomenologically guided by perception. Some children finally infer that there is a relationship between vibrations and sound but they are not able to make explicit what kind of relationship it is.

*Keywords:* Preschool education, Science education, mental representations.

## **Introducción**

El estudio de la formación representacional como construcción de conocimiento ha sido estudiado por diversos autores (Cook, 2006; Mazens y Lautrey, 2003; Osborne y Freyberg, 1998; Postigo y López-

Manjón, 2012; Pozo y Gómez-Crespo, 1998). Si bien todo conocimiento es una construcción representacional, partimos del supuesto de que algunas de estas representaciones influyen en la formación de otro tipo de representaciones durante la alfabetización científica.

En el presente trabajo utilizamos el concepto de *representaciones intuitivas* para referirnos al conjunto de ideas que poseen las personas para interpretar los fenómenos naturales y que pueden o no, estar en contradicción con lo establecido en las teorías, principios y leyes del conocimiento científico. Consideramos que las representaciones son el fruto de la percepción y estructura cognitiva del sujeto y están basadas en sus experiencias cotidianas, tanto físicas como sociales, que dan como resultado un conocimiento empírico de la ciencia (Moreira y Greca, 2005; Pozo y Gómez-Crespo, 1998).

Es así que, sin importar el nivel escolar, todos los estudiantes llegan al salón de clases con representaciones intuitivas que en muchas ocasiones son fuertes y persistentes. Por ende, estudiar cómo se forman las representaciones intuitivas es una tarea importante, pues las representaciones del mundo determinan lo que las personas hacen y pueden hacer.

En el presente trabajo se analizarán las representaciones intuitivas de niños de nivel preescolar con respecto al sonido. En el siguiente apartado se presenta una descripción sobre las representaciones, una revisión de trabajos previos relacionados con el tema, así como el marco interpretativo de la presente investigación, y finalmente se presentarán los resultados obtenidos a partir del estudio realizado y las conclusiones del trabajo.

## **¿Cómo desarrollan los niños su conocimiento del mundo?**

Comprender el proceso de aprendizaje de las personas es un tema que ha preocupado a diversas áreas de investigación, desde las neurociencias hasta las artes y humanidades a partir de las cuales se han desarrollado diversos enfoques para explicar este proceso. Si bien, el estudio del aprendizaje de las ciencias en estudiantes tiene una larga tradición, hasta hace pocos años se comenzaron a realizar investigaciones que dan cuenta de la problemática en el aprendizaje de las ciencias en niños en edad preescolar.

Los niños, para dar cuenta de cómo funciona el mundo, establecen relaciones entre las cosas y los hechos, estas relaciones les ayudan a explicarse los fenómenos naturales, las relaciones físicas entre los objetos, el papel de los otros y de ellos mismos. A este conjunto articulado de relaciones se le ha llamado *representaciones*. Estas representaciones nos permiten, a todos, explicar lo que sucede y dar sentido a los fenómenos, así como anticipar lo que va a suceder (Chi y Roscoe, 2002; Duschl y Grandy 2008; Nersessian, 1999; Tytler y Prain, 2010). Por ejemplo, una representación viso-espacial, permite situar lo que sucede en un lugar y establecer relaciones de dependencia entre acontecimientos, pero es la representación “temporal” la que da sentido a la causalidad. Disponer de una representación “adecuada”, es decir, que brinde la posibilidad de “experimentar” de forma más eficaz (no necesariamente con objetos) abre a los individuos más posibilidades de representación del mundo. Sin embargo, tener una representación del mundo no es suficiente, necesitamos hacerla explícita a nosotros mismos y a los otros de modo que genere nuevos significados para los individuos.

Además de la función explicativa, las representaciones tratan de satisfacer otras necesidades del sujeto, no necesariamente de tipo cognitivo, sino que contienen aspectos ideológicos, motivacionales o afectivos. Tampoco están completamente articuladas y disponibles en cualquier instante. Es decir, en el momento en que se demanda, el sujeto, combina distintos elementos de los que dispone para responder a la nueva demanda de representación. Los elementos de construcción están disponibles siempre, pero la articulación dentro de la estructura que le da sentido sólo se presenta para responder a una necesidad en un momento determinado (Tytler y Prain, 2010). La demanda puede ser para resolver un problema práctico o para explicar un fenómeno que acontece.

La función explicativa de las representaciones se manifiesta cada vez que el sujeto tiene que resolver un problema por lo que las representaciones están ligadas a los fines que se plantea el sujeto, es importante mencionar que esas representaciones que cada uno tiene que formar para desenvolverse en el mundo son anteriores al conocimiento científico y existían antes de que surgiera la ciencia (Delval, 2007).

Cabe señalar que uno de los problemas del estudio de las representaciones, es que no puede llegarse a ellas directamente, sino sólo de una forma indirecta, infiriéndolas a partir de lo que hacen o dicen los sujetos sobre los objetos o fenómenos (Driver, 1989; Nersessian,

1999; Tytler y Prain, 2010). Es por ello que el lenguaje y la capacidad de utilizar sistemas abstractos de representación aumentan enormemente las posibilidades de explicitación, pero estos sistemas abstractos sólo son el vehículo mediante el cual se hacen manifiestas esas representaciones.

## El papel del cuerpo en la construcción representacional

Diferentes autores han utilizado el concepto “cognición corporal” *embodiment* (Pozo, 2003) como marco interpretativo para conocer y explicar las representaciones que tienen las personas acerca de algún concepto científico y/o físico (Alibali y Nathan, 2012; Calvo y Gomila, 2008; Edwards, 2009; Gibbs, 2005; Hostetter y Alibali, 2008;). Este concepto puede analizarse desde el desarrollo de la psicología de Jean Piaget, al afirmar que las capacidades cognitivas tienen un fundamento en las habilidades sensoriomotoras (Wilson, 2002), por lo que, considera que los procesos cognitivos están fuertemente arraigados a las interacciones sensorio-motoras del cuerpo y con el mundo.

Ahora bien, ¿qué función o funciones desempeña el cuerpo en el proceso cognitivo? Según Clark y Chalmers (1998), y Wilson (2002) existen al menos dos funciones distintas pero relacionadas entre sí para comprender el papel del cuerpo en el proceso cognitivo.

La primera sugiere que el cuerpo puede funcionar como una *limitante cognitiva*: en este sentido, hablar o pensar acerca de los objetos, ya sean concretos o abstractos, implica la apelación, el despliegue o la reactivación de patrones específicos de la actividad corporal (el cuerpo da forma a la naturaleza de la actividad cognitiva y el contenido de las representaciones procesadas). Ejemplo de ello son; a) *la percepción del color*, los conceptos y experiencias de los colores, se deben a las propiedades de las células de la retina y las características del aparato visual, por lo que solo vemos el espectro de la luz visible, b) *la detección del sonido*, que debe su peculiaridad a la distancia entre los oídos, y c) *la metáfora espacial*, que es la forma en que conceptualizamos el cuerpo, a partir de las experiencias “encarnadas”, por ejemplo, la metáfora conceptual *el tiempo es un objeto movible* expresa una acción y el “lunes viene antes del martes”, hace referencia al movimiento a través del tiempo “lo que viene por delante”, “el día que quedó atrás”, entre otras (Thompson et al., 1992).

La segunda función destaca las diferentes formas en que el cuerpo actúa como un *distribuidor para el procesamiento cognitivo*, operando como un sistema de interpretación parcial de la cognición. Esto es, forma parte del proceso cognitivo, el cuerpo se “extiende” en tareas cognitivas dentro de estructuras que no se limitan al sistema nervioso central (no neuronales), funcionando como realizador parcial de procesos mentales. Algunos ejemplos de procesos no neurales son principalmente el gesto y las posturas, que se ven reflejados en estudios acerca de operaciones cognitivas, la producción del lenguaje, la plasticidad cortical y la adquisición de la habilidad motora de la mano (Rowe y Goldin-Meadow, 2009; Rowe et al., 2008).

Asimismo, otras investigaciones sugieren que el aumento del volumen de las representaciones de movimientos es modelada por las experiencias encarnadas, y que los cambios inducidos por el cuerpo ayudan a regular mejor el cerebro, con el procesamiento de la información y desarrollo cognitivo (Hlustik et al., 2004). Lo que sugiere que el cuerpo tiene un papel importante en la formación representacional.

Es así que la actividad meramente corporal, específicamente los gestos, se han considerado una forma de comunicación, para algunos autores los movimientos de hecho juegan un papel cognitivo en el desarrollo del vocabulario y el lenguaje (Rowe y Goldin-Meadow, 2009; Rowe et al., 2008). Además de considerar su trascendencia en la comprensión del significado y la formación de las representaciones.

Uno de los trabajos más citados en relación con la función del gesto en el proceso de formación de representaciones realizado por Corballis (2003), quien menciona que en este proceso participan un tipo de neuronas que transmiten la información referente a las acciones de los otros, llamadas *neuronas espejo*. Se trata de un sistema de representación análogo e icónico en su origen, subyacente a la capacidad de señalar, imitar y dirigir la vista con intención comunicativa. Por lo que la función de estas neuronas es la comunicación no verbal (señalar, mostrar, demostrar, expresar emoción, entre otras), relacionadas con la pantomima y la capacidad de imitar lo que otros hacen (mímica).

De acuerdo con McNeill (1992, 1998, 2008) es posible estudiar cuatro dimensiones del gesto: a) gesto deíctico (pointing); son los gestos para indicar objetos o localizaciones en el ambiente físico (señalar), b) gesto icónico (iconic); son los gestos que mantienen una relación perceptual de similitud entre su forma y el contenido semántico de la representación (el

modo de realización del gesto encarna aspectos retratables del contenido semántico), c) gesto metafórico; el contenido semántico se da a través de la metáfora (ej. juntar las manos como “sosteniendo” la idea, la cual sería la metáfora de la idea como objeto), y d) gesto rítmico (beat gesture); son movimientos motores y/o gestos rítmicos que no poseen contenido semántico, pero que están contenidos en la prosodia del contenido semántico. Siguiendo la lógica del trabajo de Corballis (2003) y McNeill, los gestos mímicos son altamente icónicos, debido a que constituyen un principio básico de la intención comunicativa.

## **Las representaciones intuitivas y la alfabetización científica**

Las representaciones intuitivas son construcciones personales con las que llegan los estudiantes a sus clases de ciencias, independientemente de su edad, género, cultura o nivel educativo. Muchas de estas representaciones se han formado a partir de la experiencia de los individuos con los fenómenos cotidianos y guiadas por la percepción. No son fáciles de identificar porque son parte del conocimiento implícito del sujeto, además de ser muy resistentes por lo que a menudo se convierten en fuertes barreras a la comprensión de los conceptos científicos y consecuentemente son difíciles de modificar, por lo que no se alteran por medio de la enseñanza tradicional de la ciencia. Estas representaciones no son exclusivas de los estudiantes, muchos profesores comparten las representaciones intuitivas que tienen sus alumnos, aunque no sean conscientes de ello (Carretero et al., 1996; Driver, 1989; Driver y Erickson, 1983; Pozo y Carretero, 1992; Wandersee et al., 1994).

Con respecto a la resistencia a la modificación de las representaciones intuitivas, son dos las principales explicaciones. Una de ellas supone que aquellas representaciones que están estrechamente ligadas a situaciones de la vida cotidiana y tienen un grado de coherencia y solidez variable en donde son adecuadas, son más difíciles de modificar (Pozo y Carretero, 1992). La otra es la falta de conocimiento previo, si uno no dispone de cierto nivel de conocimiento, difícilmente puede entender los argumentos presentados para conducir el cambio de esas representaciones (Strike y Posner, 1985; Chinn y Brewer, 1993; Schumacher et al., 1993).

## Representaciones intuitivas acerca del sonido

Las investigaciones que han estudiado las representaciones intuitivas relacionadas con el concepto de sonido son escasas, probablemente la razón sea que la idea de la transmisión del sonido puede ser un concepto difícil de entender.

Algunos estudios sugieren que, si bien los niños consideran al sonido como un objeto, no le otorgan todos los criterios de una entidad física. Mazens y Lautrey (2003) entrevistaron a un grupo de 89 alumnos de 6 a 10 años de edad, quienes respondieron a una serie de preguntas para determinar si confieren al sonido propiedades físicas como: sustancialidad (tiene características de sustancia), trayectoria (sigue una trayectoria lineal entre dos puntos), permanencia (viajara desde su origen en una trayectoria longitudinal infinita de tiempo) y/o peso (se ve afectada por la gravedad). Estos autores encontraron que el 46% de los niños que entrevistaron consideran que “el sonido se fue a través de agujeros” de los objetos, el 33% cree que el sonido viaja directamente a la oreja o rebota en las superficies antes de que llegue al oído de una persona, el 20% cree que el sonido “continúa para siempre” (una característica infinita de tiempo) y el 14% pensaba que el sonido tenía peso. Este estudio mostró que cuando los niños perciben el sonido como un objeto atribuyen la característica de sustancialidad más a menudo que el peso o la permanencia.

Un año después Lautrey y Mazens, (2004) analizaron la organización de las representaciones intuitivas acerca del concepto de “sonido” y las compararon con el concepto de “calor”, además de investigar el proceso de cambio conceptual que ocurre en dichas representaciones. Para ello, entrevistaron individualmente a 83 niños y niñas (aprox. 50% por género) de 8 años de edad para ver si atribuían al sonido las propiedades de los objetos (como sustancialidad, peso, permanencia y trayectoria) o las propiedades de los procesos físicos (por ejemplo, la transmisión de proximidad). Estas dos categorías darían cuenta del cambio conceptual. Sus resultados indicaron que los niños le atribuyen propiedades de objeto al sonido de manera jerárquica. Es decir, en el nivel básico (primer nivel jerárquico) le atribuyen la cualidad de permanencia, en el siguiente nivel la cualidad del peso, y, finalmente, en el tercer nivel, se le atribuye la cualidad de sustancia, es decir, el sonido se asume como un tipo de materia, lo que implicaría que no puede pasar a través de los sólidos.

Sus resultados también muestran que estas representaciones intuitivas parecen estar relativamente organizadas y el cambio conceptual es un proceso de largo plazo. Para Whittaker (2012) la física es un tema donde los alumnos tienen un gran número de representaciones intuitivas que no son compatibles con el conocimiento científico además de estar profundamente asentadas. Para este autor, el sonido es un buen ejemplo de esto, ya que requiere de la visualización de una forma de energía que se transmite a través de un medio invisible de forma invisible. Para identificar las representaciones intuitivas de los estudiantes en cuanto a la naturaleza del sonido y su transmisión, realizaron entrevistas individuales a 28 alumnos de 11 a 14 años. Algunas de las preguntas empleadas en la entrevista fueron las siguientes: 1. Su profesor está hablando en clase, ¿cómo puedes oírlo?, 2. El profesor está hablando en voz alta en el aula de al lado, ¿cómo puedes oírlo?, 3. Si sales del salón y gritas, ¿hasta dónde iría el sonido?, 4. ¿Podrías escuchar a alguien hablando junto a ti si estuvieran en el espacio? Del mismo modo que en los estudios descritos previamente (Lautrey y Mazens, 2003; Mazens y Lautrey, 2004), Whittaker concluye que las representaciones de los niños sobre el sonido se basan en la intuición sustancialista. Es decir, los alumnos le atribuyen al sonido propiedades de objeto como: trayectoria, pueden traspasar objetos y se desplaza.

Debido a que son pocos los estudios que se han ocupado de las representaciones intuitivas de los estudiantes acerca del sonido, Eshach y Schwartz (2006) apoyados con el “esquema de sustancia” nombrado y diseñado por Reiner, et al. (2000) se plantearon identificar si dicho esquema está presente en el pensamiento de los estudiantes de secundaria y examinar cómo los estudiantes utilizan las características del esquema para explicar el sonido. Además, averiguar si las características del esquema de sustancia se utilizan con cierta coherencia “local” es decir, de manera individual o si se pueden identificar consistencias más coherentes a nivel “general” entre las propiedades que los estudiantes utilizan para explicar los fenómenos sonoros. Para ello realizaron entrevistas a profundidad con preguntas abiertas y estandarizadas a diez estudiantes de secundaria. Los resultados muestran que, partir del esquema de la sustancia, el sonido fue percibido por los participantes como “algo” que se puede empujar, friccionar o controlar. Además, como una sustancia diferente a la normal con respecto a su estabilidad, naturaleza corpuscular, propiedades aditivas y características inerciales.

En otras palabras, las representaciones del sonido de los estudiantes no parecen encajar con el esquema de Reiner, et al. (2000) en todos los aspectos. Los resultados también indican que la representación de sonido de los estudiantes carece de coherencia interna, ya que conforme se les cuestionaba acerca del alcance explicativo de sus respuestas, éstas ya no eran consistentes. Con respecto a la coherencia local y general, los autores reportan que la concepción del sonido de los estudiantes coincide con la propuesta de diSessa (2002) quien considera que dichas representaciones están conectadas vagamente, como una colección fragmentada de las representaciones. Eshach y Schwartz (2006) concluyen que la noción de que el sonido se percibe sólo como “una especie de material” (cualidades de sustancia). Además de mencionar que se requiere de más investigación que ayude a entender la pertinencia del “esquema de la sustancia” de Reiner et. al (2000) que se aplica en dicha investigación.

Por último, Perales (1997) realizó un trabajo con base en las concepciones sobre la acústica en alumnos de distintos niveles educativos, partiendo de un análisis del contenido científico a tres bandas: estructural, didáctico y sociológico. Este autor concluye que, de forma análoga a como ocurre con la luz, el sonido es identificado por su causa (emisión) y por sus efectos (detección), aunque también se hace como una entidad en el espacio (propagación), por lo que representa un avance con relación a lo que ocurre con la luz. En este sentido, la evolución por edad o nivel educativo es evidente: el sonido pasa de ser un «ruido» a una propagación de ondas en el espacio material.

Perales (1997) considera que esto supone en sí mismo un avance cualitativo importante en la comprensión de los estudiantes, dado que muestra indicios de atribuir la propagación del sonido al conjunto de las partículas del medio, aunque la naturaleza de las ondas dista mucho de ser comprendida por ellos. Finalmente, sugiere que esta evolución se manifiesta en cuanto a la riqueza semántica del término *sonido* que, de ser considerado mayoritariamente como ruido por los alumnos más jóvenes, se convierte en un glosario de adjetivos relativos a sus cualidades en los de mayor edad. Asimismo, es de destacar la significativa presencia de la «música» como descriptor del sonido en los individuos de menor edad.

Los estudios descritos brindan un panorama general del tipo de representaciones que elaboran los estudiantes acerca del fenómeno del sonido. En general, resulta evidente que existe poca investigación sobre las representaciones intuitivas sobre el sonido y no hay estudios reportados con población en edad preescolar.

Tomando como base la idea de que los individuos construyen representaciones intuitivas sobre los fenómenos que observan y que el uso de los gestos (cognición corporal) son parte de esas representaciones, el objetivo de la presente investigación radica en identificar las *representaciones intuitivas* de los niños sobre el sonido y su relación con la formación del conocimiento científico en preescolar. El análisis se hace a partir de un estudio de casos, dado que implica un proceso de indagación sistemático y a profundidad que permite comprender la complejidad del problema de investigación y establecer niveles explicativos del mismo dentro de un proceso o situación dada (Stake, 1998). Además de ser inductivo, ya que las observaciones son detalladas, por lo que permiten estudiar múltiples y variados aspectos, examinarlos en relación con los otros y, al tiempo, verlos dentro de sus ambientes (Pérez Serrano, 1994).

## Método

La investigación es de corte cualitativo y corresponde a un estudio de caso.

## Muestra

Participaron seis estudiantes de 3° de preescolar en total (2 niñas y 4 niños), que tenían entre 5 y 6 años de edad y pertenecían a dos regiones distintas, dos de los niños de la comunidad de Tesigtan, Puebla y cuatro de la Ciudad de México.

## Instrumentos

La recolección de datos se obtuvo mediante la aplicación de una entrevista semiestructurada que abordó tres aspectos del sonido: producción (cómo se puede producir sonido), percepción (cómo se percibe el sonido, quién puede hacerlo, etc.) y propagación (cómo viaja el sonido, hasta dónde puede llegar, etc.). En la tabla I se muestran ejemplos de algunas de las preguntas.

TABLA I. Estructura de la entrevista

Tema	Pregunta
<b>Producción</b>	Observa a tu alrededor, ¿con cuáles cosas puedes hacer sonidos?
<b>Percepción</b>	Observa estas imágenes (perro, ballena, bebé, niño(a), gallina, árbol, piedra, pájaro, mariposa, pez, etc.) ¿quiénes pueden escuchar?, ¿por qué pueden escuchar?, ¿cómo sabes que escuchan? Si tú le hablas a la mariposa, ¿te podrá escuchar?, si tú le hablas al árbol, ¿te podrá escuchar?
<b>Propagación</b>	Se golpea el triángulo. Cuando se deja de escuchar se pregunta por dónde se fue el ruido, qué tan lejos se fue. ¿Qué pasa con ese sonido, se va para siempre, se detiene o regresa?

Fuente: Elaboración propia.

Debido a que el sonido presenta un nivel de abstracción y que es difícil de explicitar para niños tan pequeños, se utilizó material gráfico e instrumental el cual permitió contextualizar y delimitar la entrevista. También este material permitió apoyar el proceso de construcción representacional del niño al utilizarlo para ubicarlos en una situación específica en la cual pudieran producir y percibir el sonido (al tiempo que se realizaron las preguntas) permitiendo a los niños utilizarlo al explicitar sus representaciones. La tabla II describe el material utilizado durante la sesión.

TABLA II. Material utilizado durante la entrevista

Tema	Material
<b>Producción</b>	Fichas de imágenes correspondientes a personas, partes del cuerpo, instrumentos u objetos con los cuales se puede producir sonido. Fichas de imágenes correspondientes a personas u objetos los cuales pueden producir sonido por sí solos (camión, triciclo, perro árbol, ballena, pájaro, lluvia, niña(o), bebé, piedra).
<b>Percepción</b>	Fichas de imágenes correspondientes a personas y animales. Hoja de papel que se coloca frente a la cara del entrevistador, entre él y el niño. Marimba con teclas de tres materiales distintos (hule, madera y metal).
<b>Propagación</b>	Un reloj con alarma, cuatro cajas de: madera, vidrio, cartón y acrílico. Suponiendo que el reloj se coloca en cada caja. Triángulo musical metálico. Teléfono de manguera. Un par de orejeras. Un tambor y una barra metálica suspendida a centímetros del mismo.

Fuente: Elaboración propia

## Procedimiento

Previa autorización por escrito de los padres y autoridades de la escuela, las escuelas proporcionaron el espacio para realizar la entrevista. Las sesiones tuvieron una duración promedio de 30 minutos por cada niño(a), distribuidas en dos aplicaciones por día, las cuales fueron registradas en audio y video.

En principio la secuencia de la entrevista se planeó igual para todos los participantes. Sin embargo, en algunos casos se cambió el orden con la intención de no perder la atención de los entrevistados y comprender su proceso de pensamiento.

## Análisis de datos

Una vez transcritas las entrevistas se realizaron paralelamente dos análisis diferentes. En el primer análisis se construyeron categorías de análisis con el fin de organizar los datos e interpretarlos. Esta categorización de las respuestas se realizó a partir de la teoría fundada (Charmaz, 2006) que plantea que la teoría se elabora y surge de los datos analizados. Se analizaron las explicaciones y acciones que realizaron los niños (as), las cuales constituyeron las categorías de respuesta de los sujetos. En la Tabla III se describen los criterios utilizados para definir las categorías, estos criterios se construyen a partir de las ideas que los niños expresaron durante las entrevistas.

TABLA III. Categorías de análisis de la entrevista

Criterio	Categoría
Solo con letras y palabras se puede producir sonido. El sonido lo producen las letras, los demás sonidos son ruido.	1. Palabras- Sonido
Es posible producir sonido si se realiza una acción sobre un objeto concreto. Ej: Tocar, azotar, tallar, soplar, arrastrar, tirar y romper. Los sonidos son diferentes debido a que el material que lo produce es distinto.	2. Objeto- Acción- Material
Producir sonido con el cuerpo y/o percibir sonido con los oídos. Las cosas que tienen vida pueden escuchar. Sin embargo, un árbol no escucha. Por lo que se infiere que las plantas en general son consideradas objetos.	3. Cuerpo-Sonido
	4. Características de tipo sustancialista
El sonido está dentro de los objetos, por lo que es necesario sacarlo de ellos y puede ser atrapado.	4.1 Contención
El sonido puede traspasar el material que es menos duro.	4.2 Traspasa objetos
El sonido sigue un camino definido hasta los objetos u oídos. Se puede detener y continuar la trayectoria al producirse de nuevo.	4.3 Trayectoria
La "intensidad" del sonido depende de la fuerza con que se toque el objeto. Existen sonidos que pesan más que otros. Es decir, los sonidos son diferentes dependiendo del peso.	4.4 Fuerza-Peso

Fuente: Elaboración propia.

En el segundo análisis se identificaron y categorizaron los gestos utilizados por los niños para explicar sus respuestas o en algunos casos siendo el gesto la propia respuesta, siguiendo la teoría fundada. En la tabla IV se presentan los gestos identificados en la entrevista, así como su descripción.

TABLA IV. Dimensiones utilizadas para el análisis del gesto

Tipo de Gesto	Descripción
Gesto deíctico (pointing)	Señalar con el dedo índice o con toda la mano, acompañado de la respuesta verbal o no.
Gesto icónico (iconic)	Movimientos amplios (generalmente con la mano) que tienen relación con el contenido semántico de las respuestas.
Gesto metafórico	Está relacionado con el contenido semántico de la respuesta
Gesto rítmico (beat gesture)	Movimientos motores y/o gestos rítmicos que no tienen relación con el contenido semántico de la respuesta.

Fuente: Reconstruido a partir de McNeill (1992, 1998 y 2008).

## Resultados

A continuación, se presentan los resultados obtenidos a partir del análisis de las entrevistas. En la primera parte se describen las categorías identificadas en las respuestas de los niños y que corresponden a las presentadas en la Tabla III. Posteriormente se describen los gestos utilizados a lo largo de la entrevista. El análisis se hace para los tres temas abordados: percepción, producción y propagación del sonido.

La tabla V muestra las frecuencias observadas para cada una de las categorías de análisis (Tabla III) por cada uno de los temas en el total de muestra.

**TABLA V.** Frecuencias de las categorías de análisis presentes en la formación representacional de los niños respecto al sonido de acuerdo al tema

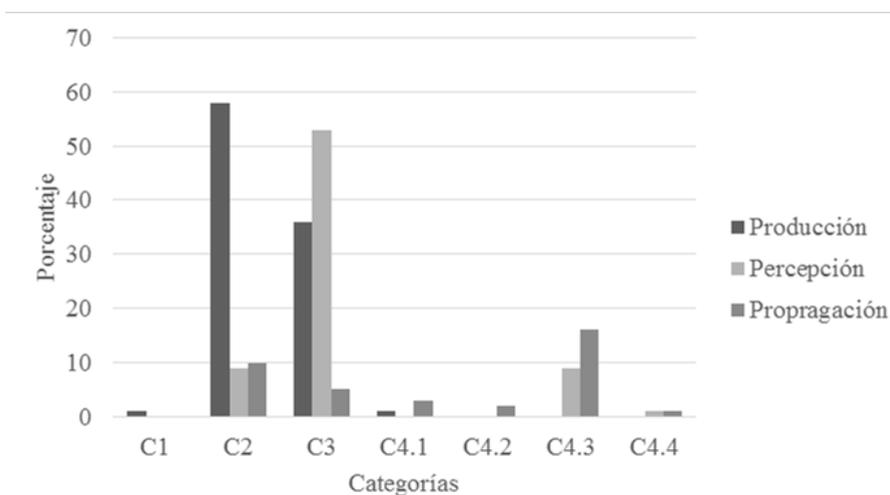
Tema	Categorías						
	C1 <i>f</i>	C2 <i>f</i>	C3 <i>f</i>	C4.1 <i>f</i>	C4.2 <i>f</i>	C4.3 <i>f</i>	C4.4 <i>f</i>
Producción	1	58	36	1	0	0	0
Percepción	0	9	53	0	0	9	1
Propagación	0	10	5	3	2	16	1

Fuente: Elaboración propia.

Nota: C1. Palabras-Sonido, C2. Objeto- Acción-Material, C3 Cuerpo-Sonido, C4.1 Contención, C4.2 Traspasa objetos, C4.3 Trayectoria-Sonido y C4.4 Fuerza-Peso.

En el gráfico I, se observan cuáles y en qué porcentajes (calculados sobre el total de las respuestas de los niños) son utilizadas estas mismas categorías para cada tema.

GRÁFICO I. Porcentaje de categorías del análisis por tema



Fuente: Elaboración propia.

De un total de 205 respuestas categorizadas (gráfica 1), se puede observar que la categoría C1 se presenta únicamente en un caso, probablemente porque el estudiante se encontraba en proceso de aprendizaje de la lecto-escritura y consideró que el sonido únicamente puede hacerse con palabras. Las categorías C2 (37.56%) y C3 (45.08%) presentan el mayor porcentaje de la representación del sonido, las cuales son utilizadas para explicar la producción y la percepción del sonido en todos los casos. La categoría C4 (16.1%), constituida por las categorías C4.1, C4.2, C4.3 y C4.4, está presente en las representaciones de todos los alumnos, y si bien, es utilizada en menor medida está más vinculada con los temas de percepción y propagación del sonido. Es necesario señalar que el caso 2 no dio suficiente información por lo que únicamente utilizó la categoría C2 y C4.3 para explicar la producción y para la propagación. Estas categorías dan cuenta de las respuestas que los niños dieron a todas las preguntas de la entrevista, sin embargo, como se mencionó anteriormente, muchos de ellos emplearon gestos para responder a las preguntas.

El segundo análisis corresponde a la identificación de los gestos que emplearon los niños para responder a las preguntas de la entrevista y que en diversas ocasiones constituyeron la respuesta en sí. En la Tabla

IV se describen las frecuencias de los gestos empleados en cada una de las categorías de análisis y para cada temática abordada. Este análisis muestra que los niños emplearon sólo dos tipos de gesto, el deíctico y el icónico.

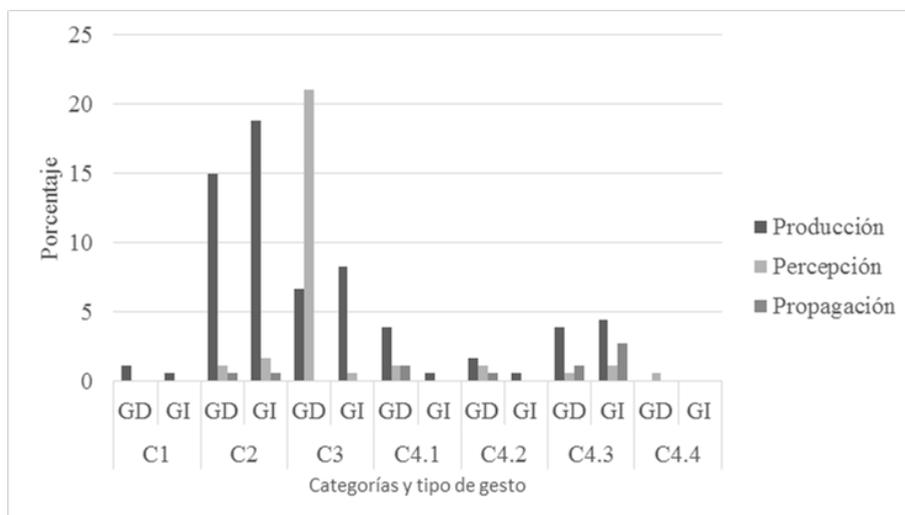
**TABLA VI.** Frecuencias de los gestos utilizados en las categorías de análisis presentes en la formación representacional de los niños respecto al sonido de acuerdo al tema

Tema	Categorías y tipos de gestos													
	C1		C2		C3		C4.1		C4.2		C4.3		C4.4	
	GD	GI	GD	GI	GD	GI	GD	GI	GD	GI	GD	GI	GD	GI
Producción	2	1	27	34	12	15	7	1	3	1	7	8		
Percepción			2	3	38	1	2		2		1	2	1	
Propagación			1	1			2		1		2	5		

Fuente: Elaboración propia.  
 Nota: GD: gesto deíctico; GI: gesto icónico

En la gráfica II se muestran los porcentajes (calculados sobre el total de los gestos registrados de la muestra) y tipo de gestos utilizados por categoría y tema.

**GRÁFICO II.** Porcentaje de uso de gestos por tema y categoría.



En el gráfico II se observa que, de los 182 gestos analizados, nuevamente las categorías C2 (37.57%) y C3 (36.46%) son las que presentan mayor frecuencia de uso. La categoría C4 en general ocupa el tercer lugar (24.31%) distribuida principalmente entre las sub categorías C4.1, C4.2 y C4.3. Por otro lado, el tipo de gesto más utilizado por los niños es el Deíctico (señalar cosas) (54.7%) seguido por el Icónico, movimientos amplios relacionados con el contenido (39.23%).

También podemos mencionar que en el tema de producción es en el que los gestos se utilizan con más frecuencia (65.19%), en segundo lugar se encuentra el tema de percepción (28.18%) y finalmente el de propagación (6.63%).

Es importante puntualizar que las otras formas de gesto identificadas en la literatura, metafórico y rítmico, no aparecieron en las expresiones de los niños en ninguna de las categorías analizadas, es probable que sea debido a que el gesto metafórico comprende una construcción más abstracta que los gestos deíctico e icónico, es decir que, en lugar de representar objetos o eventos concretos representan conceptos y relaciones. Por lo que, al considerar el sonido con características propias de un objeto, los niños aún no logran utilizar un elemento metafórico para describirlo. Y en el caso del gesto rítmico, a diferencia del gesto icónico, no es comunicativo. Se trata de representaciones que no están ligadas al contenido semántico sino a la producción oral.

El análisis de los datos revela que en todos los casos, los niños tienen una representación basada en las acciones para producir sonido (C2) y la forma en que percibe el sonido (C3), pero que también se encuentra en proceso de construcción una representación de tipo sustancialista del sonido (C4) es decir, que tiene propiedades análogas a las de los objetos (peso y trayectoria definida) presentes de la siguiente manera; las subcategorías C4.2 y C4.3 están presentes en todos los sujetos, la subcategoría C4.1 únicamente aparece en un estudiante y finalmente la subcategoría C4.4 aparece en dos de los estudiantes.

Esta idea de la sustancialización del sonido es semejante a la reportada en otras investigaciones, si bien las características de las muestras analizadas, en especial la edad, son distintas. En cuanto a la característica sustancialista otros estudios sugieren que, si bien los niños consideran el sonido como un objeto, no aplican todos los criterios de una entidad física (Mazens y Lautrey, 2003), lo que puede también observarse en los resultados de este estudio sobre el tipo de respuestas sustancialistas que

aportan los estudiantes. En las respuestas analizadas esta representación está apoyada por elementos centrados en las características de los materiales y las acciones que los estudiantes realizan para producir o percibir el sonido.

En el caso de los gestos, se observó que en general el gesto más utilizado en todos los casos para expresar sus representaciones fue el gesto *deíctico*, principalmente para el tema de producción y percepción disminuyendo su frecuencia para el tema de propagación. Esto se debe a que a nivel jerárquico es la dimensión más básica (McNeill, 1992, 1998, 2008) y puede sustituir la comunicación verbal (señalar, mostrar, demostrar, expresar emoción, entre otras). La categoría C1 la presentó el caso 1, la categoría C2 la presentaron los casos 1, 3, 4, 5 y 6. Para la categoría C4; la subcategoría C4.1 la presentan los casos 1, 3, 4, 5 y 6, la subcategoría C4.2 aparece en los casos 1, 3, 4 y 6, la subcategoría C4.3 está presente en todos los casos mientras que la subcategoría C4.3 se presenta únicamente en el caso 6.

En el caso del gesto *icónico* la frecuencia de uso fue mayor para explicar la producción del sonido correspondiente a las categorías C1 para el caso 1, C2 para todos los casos, C3 en los casos 1, 3, 4, 5 y 6. Además, si bien en todos los casos se presenta la categoría C4, la subcategoría C4.1 se encontró en los casos 1, 3, 4, 5 y 6, la subcategoría C4.2 en los casos 1, 3, 4 y 6, la subcategoría C4.3 estuvo presente en todos los casos y la subcategoría C4.4 únicamente en el caso 6.

## Discusión y Conclusión

Durante el desarrollo de las entrevistas se pudo observar en los niños la existencia de una representación del sonido que posee dos características principales. La primera que atribuye al sonido propiedades de “objeto”, que corresponde a las categorías de análisis C2 y C3, presentes en todos los casos analizados, en la cual el sonido no puede pasar a través de sólidos, por lo que puede estar “contenido” dentro de otro sólido, por ejemplo dentro de un recipiente, y para propagarse necesita de una acción específica y un espacio sin obstáculos. La segunda característica es utilizada para explicar observaciones que en un principio contradicen la representación inicial, los niños parecen construir una “representación mental sintética” (Vosniadou, 1994), la cual puede observarse en los

casos 1, 3, 4 y 6, en éstos los estudiantes representan el sonido como una sustancia pero que puede pasar a través de un sólido. Es aquí cuando al sonido se le atribuyen propiedades distintas a los sólidos ya que consideran que puede pasar a través de objetos menos duros y/o menos pesados, sin embargo, los niños no especifican de qué forma sucede esto, lo que indica que es una concepción que está en proceso de construcción.

Ambas características responden a experiencias fenomenológicas guiadas esencialmente por la percepción de los estudiantes. Las representaciones de estos niños están basadas en una intuición materialista que sugiere una estructura sustancialista como base de sus representaciones acerca del sonido. Esto sugiere que la formación de representaciones intuitivas se da a partir de la experiencia que las personas tienen con los fenómenos cotidianos, guiadas por la percepción, y que no son fáciles de identificar porque son parte del conocimiento implícito del sujeto (Carretero et al., 1996; Pozo y Carretero, 1992).

En dos casos se registró la idea de que el sonido “no se puede ver”, sin embargo, a lo largo de la entrevista se confirmó que, aunque los niños parecen distanciarse de la idea del sonido como objeto, mantienen la representación sustancialista porque le siguen otorgando propiedades que explican su movimiento aunque no sea visible para ellos. En otros dos casos, los niños mencionaron y/o percibieron la vibración durante la producción del sonido, sin embargo, no lograron articularlo en su representación. Es decir, intuyen que existe una relación entre las vibraciones y el sonido, pero aún no son capaces de hacerla explícita de forma organizada e integrarla a su representación.

Finalmente es importante destacar que, en ambos análisis de los resultados, la información es consistente en cuanto al análisis de las respuestas verbales con respecto a las respuestas gestuales. Las categorías protagonistas para la explicación sobre producción y percepción son las relacionadas con los objetos y las acciones, así como con el cuerpo (C2 y C3). Mientras que la concepción sustancialista (C4), aunque también aparece en esos temas, se vincula un poco más con la propagación.

Si bien hay que destacar que existe una limitación en cuanto al número de participantes analizados y la generalización de los resultados, resulta relevante destacar que el estudio de las representaciones en los estudiantes, sigue siendo un tema relevante. En particular resulta conveniente destacar la aportación de este estudio preliminar que atiende un tema poco

abordado con niños de esta edad y tomando en cuenta los gestos que expresan y a partir de los cuales se pueden analizar las representaciones intuitivas que construyen para explicarse los fenómenos que observan en su vida cotidiana y la comprensión que van logrando cuando, por ejemplo se desarrollan actividades específicas para que aprendan algún tema. Como se comentó previamente, las representaciones que los estudiantes construyen tienen una fuerte influencia en sus aprendizajes posteriores y muchas de estas representaciones tienen una fuerte resistencia a la modificación porque están estrechamente ligadas a situaciones de la vida cotidiana y tienen un grado de coherencia y solidez. Por ello, es conveniente continuar realizando investigaciones que ayuden a entender la construcción de representaciones y su evolución durante el desarrollo e instrucción escolar. Es decir, ampliar la muestra de trabajo no sólo en cantidad, si no en diferentes edades para comprender cómo es que el sonido pasa de ser un “ruido” a una propagación de ondas en el espacio material y cuál es la estructura base sobre la cual se construye. Además de contrastar los resultados que se generen con las investigaciones realizadas en otros países y los posibles alcances para una propuesta pedagógica de las ciencias en este nivel escolar.

## Referencias

- Alibali, Martha W. y Mitchell J. Nathan (2012), “Embodiment in mathematics teaching and learning: Evidence from learners’ and teachers’ gestures”, *Journal of the Learning Sciences*, vol. 21, núm. 2, pp. 247-286.
- Barman, Charles R, Natalie S. Barman y Julie A. Miller (1996), “Two teaching methods and students’ understanding of sound, “*School Science and Mathematics*”, vol. 96, núm. pp. 63-67.
- Calvo, Paco y Tony Gomila (Eds.). (2008), *Handbook of cognitive science: An embodied approach*, USA, Elsevier.
- Carretero, Mario, Maite Baillo y Margarita Limón (1996), *Construir y enseñar: las ciencias experimentales*, Argentina, Aique.
- Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis*. London: Sage.

- Chi, T. Michelene y Rod D. Roscoe (2002), "The processes and challenges of conceptual change", en Limón y Mason (Eds), *Reconsidering conceptual change: Issues in theory and practice*, pp. 3-27. New York, NY: Springer.
- Chinn, Clark y William Brewer (1993), "The role of anomalous data in knowledge acquisition: A theoretical framework and implications for science instruction", *Review of educational research*, vol. 63, núm. 1, pp. 1-49.
- Clark, Andy y David Chalmers (1998), "The extended mind", *Analysis*, vol. 58, núm. 1, pp. 7-19.
- Clement, John, David E. Brown y Aletta Zietsman (1989), "Not all preconceptions are misconceptions: finding 'anchoring conceptions' for grounding instruction on students' intuitions." *International journal of science education*, vol. 11, núm. 5, pp. 554-565.
- Cook, Michelle Patrick (2006), "Visual representations in science education: The influence of prior knowledge and cognitive load theory on instructional design principles", *Science education*, vol. 90, núm. 6, pp. 1073-1091.
- Corballis, Michael C. (2003), "From mouth to hand: gesture, speech, and the evolution of right-handedness", *Behavioral and Brain Sciences*, vol. 26.02, pp. 199-208.
- Delval, Juan (2007), "Aspectos de la construcción del conocimiento sobre la sociedad (Aspects of the construction of knowledge about society)", *Educar*, vol. 30, pp. 45-64.
- diSessa, Andrea (2002), "Why conceptual ecology is a good idea", en Limon, y Mason (Eds.), *Reconsidering conceptual change: Issues in theory and practice*, New York, NY: Springer, pp. 28-60.
- Driver, Rosalind (1989), "Students' conceptions and the learning of science", *International Journal of Science Education*, vol. 11, núm. 5, pp.481-490.
- Driver, Rosalind y Gaalen Erickson (1983), "Theories-in-action: Some theoretical and empirical issues in the study of students' conceptual frameworks in science", vol. 10, pp. 37-60.
- Duschl, Richard A. y Richard E. Grandy (2008), *Teaching scientific inquiry: Recommendations for research and implementation*. USA, Sense Publishers.
- Edwards, Laurie D. (2009), "Gestures and conceptual integration in mathematical talk", *Educational Studies in Mathematics*, vol. 70, no 2, pp. 127-141.

- Eshach, Haim y Judah L. Schwartz (2006), "Sound Stuff? Naïve materialism in middle-school students' conceptions of sound", *International Journal of Science Education*, vol. 28, núm. 7, pp. 733-764.
- Gibbs Jr, Rimond, W. (2005), *Embodiment and cognitive science*, Cambridge, Cambridge University Press.
- Hluštík, Petr, Solodkin, Ana, Noll, Douglas C. y Small, Steven (2004), "Cortical plasticity during three-week motor skill learning" *Journal of Clinical Neurophysiology*, vol. 21, pp. 1-12.
- Hostetter, Autumn B. y Martha W. Alibali (2008), "Visible embodiment: Gestures as simulated action", *Psychonomic Bulletin Review*, vol. 15, núm. 3, pp. 495-514.
- Lautrey Jacques y Mazens, Karine (2004), "Is children's naive knowledge consistent? A comparison of the concepts of sound and heat", *Learning and Instruction*, vol. 14, núm. 4, pp. 399-423.
- Mazens, Karine y Lautrey Jacques (2003), "Conceptual change in physics: children's naïve representations of sound", *Cognitive Development*, vol. 18, pp. 159-176.
- McNeill, David (1992), *Hand and mind: What gestures reveal about thought*, USA, University of Chicago press.
- McNeill, David (1998), "Speech and gesture integration", *New Directions for Child and Adolescent Development*, vol. 79, pp. 11-27.
- McNeill, David (2008), *Gesture and thought*, USA, University of Chicago Press.
- Moreira, Marco Antonio, e Ileana María Greca (2005), "Cambio conceptual: análisis crítico y propuestas a la luz de la teoría del aprendizaje significativo", *Ciência & Educação. Bauru*, vol. 9, núm. 2, pp. 301-315.
- Nersessian, Nancy (1999). Model-based reasoning in conceptual change, en Magnani, L. Nersessian, Nancy, Thagard, Paul (eds.) *Model-based reasoning in scientific discovery*, USA, Springer, pp. 5-22.
- Osborne, Roger y Peter Freyberg (1998), *El Aprendizaje de las ciencias: implicaciones de las "ideas previas" de los alumnos*, vol. 121, Madrid, España, Narcea Ediciones
- Perales Palacios, Francisco Javier (1997), Escuchando el sonido: concepciones sobre acústica en alumnos de distintos niveles educativos, *Enseñanza de las Ciencias*, vol. 15, núm. 2, pp. 233-247.
- Pérez Serrano, Gloria (1994), Investigación cualitativa: retos e interrogantes. Técnicas y análisis de datos, México, La Muralla.

- Piaget Jean e Barbel Inhelder ([1966] 2008), *The psychology of the child*, New York, Basic books.
- Postigo, Yolanda y Asunción López-Manjón (2012), “Students’ Conceptions of Biological Images as Representational Devices”, *Revista Colombiana de Psicología*, vol. 21, núm. 2, pp. 265-284.
- Pozo Juan Ignacio y Mario Carretero (1992), “Causal theories, reasoning strategies, and conflict resolution by experts and novices in Newtonian mechanics” en Andreas Demetriou, Anastasia Efklides, Michael Shayer (eds) *Neo-Piagetian Theories of Cognitive Development. Implication and Applications for Education*, Londres: Routledge, pp. 231-55.
- Pozo, Juan Ignacio (2003), *Adquisición de conocimiento: cuando la carne se hace verbo*, Madrid, Morata.
- Pozo, Juan Ignacio y Miguel Ángel Gómez-Crespo (1998), “Aprender y enseñar ciencia: del conocimiento cotidiano al conocimiento científico”, Madrid, Morata.
- Reiner, Miriam, James Slotta, Michelene T. H. Chi y Lauren B. Resnick (2000), “Naive physics reasoning: A commitment to substance-based conceptions”, *Cognition and instruction*, vol. 18, núm. 1, pp. 1-34.
- Rowe, Meredith L, Şeyda Özçalışkan y Susan Goldin-Meadow (2008), “Learning words by hand: Gesture’s role in predicting vocabulary development”, *First language*, vol. 28, núm. 2, pp. 182-199.
- Rowe, Meredith L. y Susan Goldin-Meadow (2009), “Differences in early gesture explain SES disparities in child vocabulary size at school entry”, *Science*, vol. 323, pp. 951–953.
- Schumacher, G. M., Tice, S., Wen Loi, P., Stein, S., Joyner, C., y Jolton, J. (1993), “Difficult to change knowledge. Explanations and interventions. En *Proceedings of the Third International Seminar on Misconceptions and Educational Strategies in Science and Mathematics*”, Ithaca: Cornell University, en: [https://www.google.com.mx/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0ahUK Ewi5pbvxiZPSAhXJsFQKHTNWAOAQFggZMAA&url=http%3A%2F%2Fwww.mlrg.org%2Fproc3pdfs%2FSchumacher\\_Knowledge.pdf&usg=AFQjCNFXI-ZtakrFEGrGN6-f2s1lZR55Bw&sig2=V3BD2n14sC2Rnp5zEmzw-w](https://www.google.com.mx/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0ahUK Ewi5pbvxiZPSAhXJsFQKHTNWAOAQFggZMAA&url=http%3A%2F%2Fwww.mlrg.org%2Fproc3pdfs%2FSchumacher_Knowledge.pdf&usg=AFQjCNFXI-ZtakrFEGrGN6-f2s1lZR55Bw&sig2=V3BD2n14sC2Rnp5zEmzw-w), (consulta febrero 4, 2017).
- Stake, Robert E (1998), *Investigación con estudio de casos*. Madrid, Ediciones Morata.
- Strike, Kenneth A. y George J. Posner (1985), *A conceptual change view of learning and understanding*, USA, Academic Press.

- Thompson, Evan, Adrian Palacios y Francisco Javier Varela (1992), "On the ways to color", *Behavioral and Brain Sciences*, vol. 15, núm. 1, pp. 56-74.
- Tytler, Russell y Vaughan Prain (2010), "A framework for re-thinking learning in science from recent cognitive science perspectives" *International Journal of Science Education*, vol. 32, núm. 15, pp. 2055-2078.
- Wandersee, James, Joel Mintzes y Joseph Novak (1994), "Research in alternative conceptions in science", en D. Gabel (ed.), *Research Handbook on Research on Science, Teaching and Learning* New York, N.Y.: McMillan Pub, pp. 177-210.
- Whittaker, Andrew (2012), "Pupils Think Sound Has Substance--Well, Sort of", *School Science Review*, vol. 346, pp. 109-111.
- Wilson, Margaret (2002), "Six views of embodied cognition", *Psychonomic bulletin y review*, vol. 9, núm. 4, pp. 625-636.

**Información de contacto:** Mtra. Claudia-Elena Velázquez-Olmedo. Universidad Nacional Autónoma de México. Instituto de Ciencias Aplicadas y Tecnología. Grupo de Cognición y Didáctica de las Ciencias. ICAT, Avenida Universidad N°3000, Ciudad Universitaria. Delegación Coyoacán, Ciudad de México, CP. 04510. E-mail: medeagnes@gmail.com



# Vinculación de la prueba de comprensión oral del examen CertAcles de la Universidad de Granada con el MCER<sup>1</sup>

## Linking the University of Granada CertAcles listening test to the CEFR

DOI: 10.4438/1988-592X-RE-2017-381-380

Caroline Shackleton

Universidad de Granada

### Resumen

Este estudio, como parte del proceso de validación fehaciente del uso e interpretación de las puntuaciones del examen de inglés binivel de la Universidad de Granada (UGR), pretende relacionar, oficialmente, la prueba de comprensión oral de este examen con el Marco Común Europeo de Referencia (MCER). Con tal fin seguimos las recomendaciones del Consejo de Europa (2009) para determinar unos puntos de corte representativos de los niveles B1 y B2 del MCER. La prueba utilizada en el estudio fue la convocatoria de marzo del 2017 (N=464), que demostró ser un instrumento de medición fiable y apropiado. En el estudio participaron diez expertos, quienes, tras una sesión para familiarizarse con el MCER, realizaron varias sesiones de fijación de puntos de corte. Se utilizaron dos métodos: el método *Basket*, o de la cesta, para familiarizar a los participantes con los ítems de la prueba. después del cual se analizaron los resultados empleando un modelo de múltiples facetas de Rasch (MFR); por otro lado, el método *Bookmark*, o del marcador, que incorporó parámetros de dificultad del ítem, resultantes de un análisis Rasch de los puntajes de la prueba. La mediana final de los puntos de corte aportados por los expertos se calculó empleando una probabilidad de respuesta de 0.67, y se transfirió a la escala de capacidad zeta del modelo Rasch para determinar las puntuaciones brutas que representarán

---

<sup>(1)</sup> Agradecimientos: La autora quiere agradecer al Centro de Lenguas Modernas de la Universidad de Granada por sus aportaciones en apoyo de este estudio, así como a los dos revisores anónimos por sus valiosos comentarios y sugerencias.

los dos niveles de dominio del MCER en cuestión. También se realizaron otros controles de fiabilidad y de validez para proporcionar una documentación transparente sobre el proceso completo, recomendado por el Consejo de Europa (2009). El método de equiparación de puntuaciones del modelo Rasch permite que los puntajes sean reproducibles en futuras versiones de la prueba y, por tanto, apoyan la *inferencia de generalización* del argumento de validez para la prueba binivel de la UGR.

*Palabras Clave:* pruebas de competencia lingüística, fijación de puntos de corte, validez, la teoría de respuesta al ítem

### **Abstract**

As part of the ongoing collection of validity evidence to support the interpretation and use of the UGR bi-level English test scores this study aims to formally map the oral comprehension part of the test to the Common European Framework of Reference (CEFR). The advice provided by the Council of Europe (2009) was followed in order to determine cut scores which are representative of CEFR B1 and CEFR B2 levels. The test used in the study was from the March 2017 administration (N=464) which was shown to be a reliable and appropriate measurement instrument. The study involved ten participant judges who, after a CEFR familiarisation session, took part in standard setting sessions. Two methods of standard setting were employed. The *Basket* method was used mainly in order to familiarise the participants with the test items and was analysed using a Many-Facet Rasch measurement model. The *Bookmark* method allowed for the incorporation of item difficulty parameters produced by a Rasch analysis of test scores. The final median cut scores of the judges using a response probability of .67 were mapped back to the Rasch theta ability scale in order to determine the raw scores which represent the two CEFR ability levels. Several other validity and reliability checks were also carried out following CoE (2009) and transparent documentation on the whole process is provided. By using Rasch common item equating, the scores are reproducible on future versions of the test and so support the *generalisation inference* of the UGR test's validity argument.

*Keywords:* proficiency tests (language), standard setting, validity, item response theory

## **Introducción**

En los últimos años, como parte del proceso de adaptación al Plan Bolonia y a la puesta en marcha del Espacio Europeo de Educación Superior, el Ministerio de Educación, Cultura y Deportes de España ha fomentado

una política que resalta la comunicación en una segunda lengua como competencia fundamental para aumentar la movilidad universitaria y la empleabilidad. A tal fin, es de suma importancia que se proceda al diseño y planificación de cualquier examen de carácter nacional de manera que permita la vinculación de la actuación del candidato con los niveles de competencia del Marco Europeo Común de Referencia para las lenguas (MECR) si deseamos respetar los principios de transparencia, comparabilidad y coherencia. Dichos principios juegan, sin duda, un papel de máxima repercusión para el fomento de la movilidad educacional y profesional dentro de la Unión Europea en cumplimiento de las actuales políticas de plurilingüismo europeas. En efecto, las recomendaciones del Consejo de Ministros sobre el empleo del MECR incluyen la exhortación a que los países miembros:

... aseguren que todos aquellos exámenes, pruebas, o procedimientos de evaluación que conduzcan a la obtención de títulos de lengua de oficial reconocimiento tomen en debida consideración los aspectos relevantes del uso de la lengua, así como las competencias lingüísticas expuestas en el MECR, que se lleven a cabo en concordancia con principios de buenas prácticas y gestiones de calidad internacionalmente reconocidos, y que los procedimientos para vincular dichos exámenes y pruebas a los niveles de referencia común (A1–C2) del MCER se realicen de manera fiable y transparente. (Consejo de Europa, 2008, p.4)

Como respuesta a este llamamiento, un gran número de los centros de lenguas universitarios en España han proporcionado exámenes de acreditación cuya elaboración sigue pautas internacionales, para así poder producir pruebas que sean válidas y de fiable vinculación al MCER. Además, en un intento de agilizar la coordinación y regulación de dichos esfuerzos, la Asociación de Centros de Lenguas en la Enseñanza Superior (ACLES) ha establecido un procedimiento de carácter nacional para certificaciones vinculadas al MCER: los exámenes *CertAcles*. Dichos exámenes han sido aprobados por la Conferencia de Rectores de las Universidades Españolas (CRUE) a nivel nacional, y además gozan de reconocimiento a nivel internacional a través de la Confederación Europea de Centros de Lenguas en la Enseñanza Superior (Cercles) desde el 2014. El modelo *CertAcles* (<http://www.acles.es/es>) requiere que cada destreza de capacidad lingüística sea evaluada de manera individual, que

sea desarrollada a partir de especificaciones basadas en el MCER, y que haya información proporcionada sobre los controles de calidad que siga las pautas establecidas por *los estándares* (AERA, APA y NCME, 2014).

Desde el 2009, el Centro de Lenguas Modernas (CLM) de la Universidad de Granada (UGR) administra tres veces por año un examen *CertAcles* binivel de B1/B2, de manera gratuita para los universitarios de la UGR. Desarrollado con la colaboración de la Dra. Rita Green, miembro experto de EALTA, el examen ya constituye un sistema de evaluación estable, y sigue todas las comprobaciones de validez recomendadas por ACLES. En el presente estudio describimos y reseñamos los procedimientos realizados para la vinculación del examen al MCER y el establecimiento de puntos de corte actualmente empleados para la prueba de comprensión oral del examen binivel B1/B2 de la UGR.

## Marco teórico

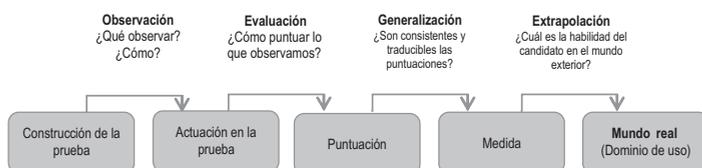
El MECR, con sus seis niveles descriptivos (A1 hasta C2), se ha convertido en el sistema de referencia de los niveles de capacidad lingüística no solo en Europa, sino en otras muchas partes del mundo (Deygers, Zeidler, Vilcu, y Carlsen, 2017; Figueras, 2012). La vinculación de los resultados de una prueba con niveles del MCER amplía el nivel significativo de los resultados de los usuarios de dichas pruebas (Kane, 2012). Por ejemplo, cuando un resultado informa sobre la capacidad de comprensión oral a nivel B1, el usuario destinatario de los resultados de la prueba recibe información sobre los tipos de actividades que un usuario de la lengua con estas competencias sería capaz de realizar (Tannenbaum y Cho, 2014). En este sentido, cualquier prueba que afirme medir las competencias del MCER, tiene la obligación de fundamentar dichas alegaciones en pruebas de que esto es así. En otras palabras, una prueba debe demostrar su validez.

La validez no se percibe como propiedad intrínseca de la prueba —una prueba no se puede validar como tal (Chapelle, 2012; Cizek, 2016)— sino más bien como un concepto polifacético y unitario que se basa en múltiples fuentes de evidencia para comprobar las inferencias sacadas sobre la actuación de un candidato y así justificar las decisiones resultantes tomadas en base a los resultados de pruebas (Messick, 1989). Es necesaria la recopilación de evidencias que respalde la interpretación

y uso de resultados de pruebas; por tanto, es recomendable el empleo de un enfoque *basado en argumentos de validez*, el cual proporciona un marco que sirve de guía para los organismos de evaluación sobre los tipos de evidencia que deben recoger (AERA, APA and NCME, 2014; Bachman, 2005; Bachman y Palmer, 2010; Chapelle, 2012; Kane, 2012).

Las decisiones sobre qué tipos de evidencia deben ser recopilados se deben basar necesariamente en la finalidad de la prueba (Fulcher y Owen, 2016) y, en el contexto europeo, el marco a seguir ha sido proporcionado por ALTE (2011) con el fin de guiar a los organismos de evaluación en el diseño y desarrollo de pruebas vinculadas al MCER. Dicho documento define, según líneas generales, los tipos de evidencias que deben ser recopiladas a lo largo del ciclo de desarrollo de la prueba (véase figura1), con el objetivo de incorporar la validación de pruebas a la totalidad del proceso de diseño e implantación de la prueba. A diferencia de las recomendaciones de Bachman (2005), y Bachman y Palmer (2010), la cadena de razonamiento comienza con el constructo de la prueba, no en decisiones basadas en los resultados de prueba. Como tal, vincula la actuaciones en las tareas de la prueba a inferencias sobre las capacidades lingüísticas del candidato en situaciones del mundo exterior, esto es, la inferencia de extrapolación. No obstante, el mundo más allá de la prueba permanece sin especificarse; se trata de una afirmación general con vínculos al MCER. Por tanto, se puede argumentar que el modelo ALTE dirige nuestro enfoque hacia la interpretación de los resultados, es decir, que pregunta ‘¿se puede interpretar un resultado de prueba como un nivel de capacidad del MCER?’ Tal postura se alinea con la conceptualización de la validez propuesta por Cizek (2016), cuyo marco diferencia entre la validación de inferencias de resultados de pruebas por un lado, y las justificaciones para la implantación de la prueba por el otro.

FIGURA I. Cadena de razonamiento para un argumento de validez



Fuente: ALTE (2011, p. 15)

Para un enfoque basado en el argumento, la evidencia de la validez del establecimiento de los puntos de corte es un componente esencial (Papageorgiou y Tannenbaum, 2016, p.110). Debido a que los candidatos tienen que alcanzar el punto de corte de una prueba para demostrar que han conseguido el nivel de capacidad requerido, los resultados de esta quedarían claramente expuestos a la crítica en el caso de que dicho punto no se estableciese de manera apropiada. A tal fin, el Consejo de Europa (2009) ha proporcionado un manual que describe, en líneas generales, el proceso necesario para que los organismos de evaluación vinculen sus pruebas al MCER. Dicho proceso de vinculación consta de las siguientes cuatro actividades interrelacionadas, todas ellas necesarias para poder relacionar los resultados de prueba con el MCER.

- **Familiarización:** Los miembros de cualquier panel de examinadores deben estar familiarizados con el contenido del MCER y sus escalas.
- **Especificaciones:** Estas deben incluir una descripción detallada de la prueba y su relación con las categorías del MCER para así construir un argumento de vinculación sobre la relevancia de esta para el MCER.
- **Estandarización y Establecimiento de normas:** Toda prueba debe estar provista de un punto de corte u otra puntuación mínima como requisito necesario para aprobarla.
- **Validación empírica:** El Consejo de Europa (2009) puntualiza tres formas de evidencia de validez —la de procedimiento, la interna y la externa— las cuales tienen que proporcionarse como parte del proceso de establecimiento de normas.

Son muchos los organismos de evaluación que han seguido los consejos del manual desde su publicación: la mayoría de los proveedores de exámenes de dominio internacionales para la lengua inglesa han realizado estudios de validación (p.ej. véase Brunfaut & Harding, 2014; Kanistra & Harsch, 2017; Tannenbaum & Wylie, 2008), y proveedores menores a) nivel nacional también siguen los pasos descritos en el manual para poder reivindicar su vinculación al MCER (p.ej. véase Downey & Kolias, 2010; Kantarcioglu, Thomas, O'Dwyer, & O'Sullivan, 2010). El enfoque principal de dichos estudios es el proceso de establecimiento de normas y su validación empírica. Como viene expuesto en el Consejo de Europa (2009), el momento decisivo en el proceso de vincular una

prueba a cualquier nivel del MCER es el establecimiento de una norma de decisión para poder asignar alumnos a uno de los niveles del MCER en función de su actuación en la prueba (p.11); es decir, la determinación del punto de corte.

En resumen, cualquier prueba que afirme tener vinculación con el MCER debe validarse y debe demostrarse que es fiable y representativa del nivel de competencia MCER en cuestión, ya que si no se demuestra empíricamente que una prueba es válida y fiable, su vinculación al MCER carece de sentido (Alderson, 2012). Por consiguiente, durante el ciclo de desarrollo de una prueba, el proceso de estandarización se suele realizar con posterioridad. Son precisamente estos tipos de evidencia de validez los que se han utilizado en la prueba de la UGR; el enfoque del presente estudio pretende describir el proceso oficial de vinculación expuesto en el Consejo de Europa (2009) y su empleo en la prueba de la UGR para establecer puntos de corte a niveles B1 y B2.

## Metodología

Los diferentes estudios de investigación clasifican las técnicas de establecimiento de normas según dos principales tipologías: la primera se enfoca en la prueba; la segunda, en el candidato. En el presente estudio, la carencia de datos sobre la población examinada requirió que el estudio principal adoptara un enfoque en torno a la prueba. No obstante, los datos tomados de un número reducido de candidatos previamente matriculados en un curso de familiarización del examen en el CLM fueron utilizados como punto de comparación para así proporcionar cierta evidencia triangulada y por tanto, contribuir a la evidencia de la validez externa.

De los diferentes métodos que tienen un enfoque en torno a la prueba, se descartaron los de tipo de probabilidad *Angoff*: no solo llevan mucho tiempo, sino que estudios previos además han señalado problemas respecto a la incapacidad de los expertos para entender y articular correctamente las probabilidades condicionales (Ferrara y Lewis, 2012; Hambleton y Jirka, 2006; Reckase, 2010). En su lugar, se decidió utilizar una combinación de dos otros métodos que actualmente tienen una gran popularidad en diferentes investigaciones: El método *Basket* y el método *Bookmark* (que en español corresponden a ‘de cesta’ y ‘marcador’, respectivamente).

A efectos prácticos, el método *Basket* es un *método de emparejamiento de descriptores de ítems*, el cual especifica el nivel de competencia necesario para que se conteste cada ítem de la prueba. Los expertos analizan los ítems de la prueba para responder a la pregunta ‘¿En qué franja del MCER tendría que estar un candidato para poder contestar este ítem de forma correcta?’. Se considera el más práctico y sencillo de todos los métodos de establecimiento de puntos de corte y, además, un método que no solo refleja la importancia de los descriptores de los niveles de competencia, sino que también enfatiza el contenido de la prueba. Sin embargo, no proporciona ninguna información sobre la dificultad de los ítems y, como tal, uno de sus mayores problemas es la falta de consistencia a la hora de comparar los juicios con las medidas de dificultad empíricas (Kaftandjieva, 2010). A pesar de esta desventaja, el método ha sido empleado en el presente estudio no para llegar a establecer el punto de corte final, sino principalmente como medio de familiarización con el contenido y para fomentar una mayor discusión sobre ello entre los expertos. Los resultados fueron analizados empleando el modelo de medición de múltiples facetas de Rasch (MMFR) en programa FACETS (Linacre, 2017). Aunque se ha argumentado que el análisis mediante MMFR no es del todo apropiado para las fases posteriores del proceso de establecimiento de puntos de corte, cuando los expertos deben esforzarse para llegar a un consenso (Eckes, 2009), dicha crítica no se consideró relevante para el contexto del presente estudio, debido a que su empleo se limitaba exclusivamente a la fase de familiarización inicial.

Para que el proceso de establecer puntos de corte tenga un valor verdaderamente significativo, los expertos deben recibir información estadística sobre el funcionamiento de los ítems de la prueba. En el método *Bookmark* (Mitzel, Lewis, Patz, y Green, 2001), los ítems son mostrados en un cuadernillo donde aparecen ordenados del más fácil al más difícil distribuyéndose, por tanto, según parámetros de dificultad. Los expertos avanzan por el cuaderno de ítems ordenados (CIO) desde el ítem más fácil hasta el más difícil, colocando un ‘marcador’ en el punto exacto en el que consideran que un candidato que cumple los requisitos mínimos tendría menos probabilidades de dar la respuesta correcta que la mínima probabilidad especificada. Se concibe la capacidad del candidato para dominar un ítem en términos probabilísticos, o ‘*probabilidad de respuesta*’ (PR); esta debe establecerse de antemano, aunque en la práctica se suele establecer en 0.67, una probabilidad del 67 % —es decir, 2/3— de que

el candidato conteste el ítem de manera correcta. El marcador se coloca entre dos ítems en la escala latente representada por los parámetros de dificultad de *Rasch* (logits) para los ítems ( $\beta$ -parameters). Finalmente, se toma la mediana de los marcadores de los expertos en la variable latente para así llegar a un estándar en común. En resumen, el método *Bookmark* no solo es fácil de implantar en pruebas que se desarrollan utilizando el modelo de medición *Rasch*, y donde la capacidad del candidato y la dificultad del ítem se colocan en la misma escala, sino que refleja de forma clara la naturaleza continua de las escalas del MCER; dos ventajas que le confieren idoneidad para su uso en el presente estudio.

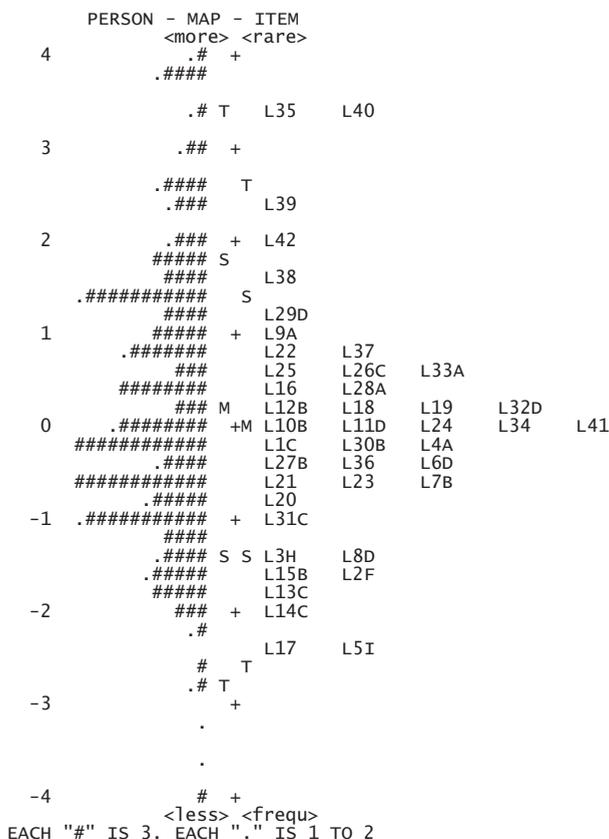
## Instrumentos

La prueba empleada para el presente estudio proviene de la convocatoria de marzo de 2017. La prueba de comprensión (42 ítems) comprende cinco tareas: dos a nivel B1, una en el umbral entre B1 y B2, y dos a nivel B2. Todas las tareas fueron desarrolladas de acuerdo con las especificaciones de la prueba y con la colaboración de nuestra asesora; por tanto, se puede considerar que su nivel de dificultad se ha establecido de manera correcta por expertos. Además, todas habían sido pilotadas de antemano ( $N =$  entre 144 y 434), con la subsiguiente revisión o eliminación de los ítems que resultaron no funcionar de forma correcta. La prueba presenta un buen muestreo del constructo basado en los descriptores de niveles B1 y B2, e incluye una variedad de métodos de evaluación/tipos de tarea.

El análisis de la prueba con teoría clásica de los tests ( $N = 464$ ) produjo un Alfa de Cronbach de 0.934 ( $EEM = 2.5$ ) y evidencia la fiabilidad de la prueba. El análisis *Rasch* de los resultados de la prueba demuestra que la prueba tiene unidimensionalidad, y las estadísticas *Rasch* de *infit mean square (MNSQ)* quedan entre los valores satisfactorios de 0.74 y 1.24. La separación de ítems es de 10.2 y la fiabilidad de la separación es de 0.99. Por consiguiente, se puede afirmar que la prueba incluye un amplio abanico de diferentes dificultades de los ítems, del mismo modo que podemos confiar en que las estimaciones de dificultad sean reproducibles. El mapa de la variabilidad de los ítems reproducido en la figura II muestra los ítems ordenados según su dificultad, junto con la capacidades de los candidatos. Utilizamos esta información para crear el CIO para el estudio que usó el método marcador. Aquí, la media de los

valores *logit* de las dificultades de los ítems y la de las capacidades de los candidatos resultaron muy parecidas, lo cual demuestra que la dificultad de la prueba se ajusta bien a la población evaluada. Dado que el presente estudio no pretende investigar la prueba en sí, consideramos que la susodicha información es evidencia suficiente de que el instrumento de medición utilizado en el estudio es tanto fiable como pertinente.

FIGURA II. Mapa de variabilidad de ítems



Fuente: Creada por la autora

De acuerdo con Cizek (2012), se les administró un cuestionario a los expertos tras las sesiones de establecimiento de puntos de corte para así suministrar evidencia de la validez de procedimiento. El cuestionario

empleaba una escala *Likert* de cuatro puntos para eliminar la posibilidad de respuestas neutras, y permitió la recopilación de las opiniones de los expertos respecto a sus interpretaciones del proceso global, además de sus pensamientos sobre las decisiones finales tomadas con respecto a los puntos de corte.

## Participantes

Resulta importante que los participantes sean expertos tanto en la aplicación del MCER, y por tanto en sus descriptores de niveles de competencia, como en la enseñanza y evaluación de la lengua evaluada (Tannenbaum y Cho, 2014). Este es, en efecto, el caso de los diez expertos que participaron en el estudio. Como la tabla I muestra, todos son profesores en el CLM con buena experiencia de la enseñanza del inglés como segunda lengua, y con buenos conocimientos del MCER. Además, conocen bien a la población de candidatos, y son conscientes de las consecuencias que los puntos de corte tienen para dicha población.

TABLA I. Información biográfica de los expertos

Experto	Sexo	Títulos	Años enseñando inglés	Número de horas de formación recibida en relación con el MCER
1	Mujer	Licenciatura, Master	35	300+
2	Hombre	Licenciatura	28	80
3	Hombre	Licenciatura, Diploma de posgrado TEFL	35	100+
4	Hombre	Licenciatura, Master, Diploma de posgrado CELTA	16	70
5	Hombre	Licenciatura	22	70
6	Mujer	Licenciatura, Master	36	400+
7	Mujer	Licenciatura, Posgrado PGCE, PhD	28	750+
8	Hombre	Licenciatura, Master, Diploma de posgrado CELTA	18	500+
9	Mujer	Licenciatura, PGCE, Master	28	750+
10	Mujer	Licenciatura	22	500+

Fuente: Creada por la autora

## Resultados

A continuación, se procederá a informar sobre los resultados de cada fase de las sesiones de establecimiento de puntos de corte, incluidas las fases de familiarización y de validación externa.

### Familiarización

Como preparación para el estudio, se les pidió a todos los expertos que realizaran sesiones de formación para la destreza de comprensión oral que han sido proporcionadas por el proyecto *Ceftrain* (véase <http://www.helsinki.fi/project/ceftrain>). Se procedió a realizar un ejercicio de ordenar descriptores del MCER para la destreza de comprensión oral, en el cual los expertos recibieron una mezcla de descriptores y tenían que asignar cada uno a un nivel del MCER. Los resultados de este ejercicio dieron un Alfa de Cronbach de 0.98. La tabla II muestra la correlación entre los descriptores del MCER y las estimaciones de los participantes obtenida mediante el empleo de un coeficiente de correlación de Spearman. Estos resultados fueron transmitidos a los participantes, para que se pudieran debatir más a fondo con el fin de llegar a un consenso sobre las características significativas de los niveles del MCER.

TABLA II. Correlación entre descriptores del MCER y asignaciones de los expertos

Ex- perto	1	2	3	4	5	6	7	8	9	10
	.851	.875	.835	.889	.896	.836	.957	.985	1.00	.870

(Nota: todas las correlaciones fueron significativas:  $p \leq .01$ )

Fuente: Creada por la autora

### El método *Basket* de establecimiento de puntos de corte

Los expertos recibieron una sesión informativa inicial sobre el constructo del examen y las especificaciones de la prueba, seguida por una explicación específica del método *Basket*. La figura III muestra un

análisis MFR de los resultados de todos los expertos para cada ítem de la prueba. El análisis permite tomar en cuenta múltiples aspectos de las estimaciones, y calibra ítems, expertos y la escala de calificación usando la misma escala de intervalos iguales. La escala empleada para representar el nivel mínimo del MCER que un candidato debería poseer para poder contestar a cada ítem es la siguiente: 1=A2, 2=B1 mínimo, 3= B1 holgado, 4=B2 mínimo, 5=B2 holgado, 6=C1. Los resultados de las dos rondas difieren mínimamente, y demuestran que los expertos confiaban en sus decisiones.

FIGURA III. Regla vertical de los resultados del método *Basket* en FACETS

Measr		+rater						+item		-round		Scale	
+	8	+						+	*	+		+	(6)
+	7	+						+		+		+	---
+	6	+						+	*	+		+	
+	5	+						+	**	+		+	5
+	4	+						+	*	+		+	---
+	3	+						+	***	+		+	
+	2	+						+	*	+		+	4
+	1	+	4					+	*	+	1	2	+
	0	+	1	2	3	7	8	+	*	+		+	---
			5	6	9								
+	-1	+	10					+	*	+		+	3
+	-2	+						+	*	+		+	
+	-3	+						+	**	+		+	---
+	-4	+						+	***	+		+	
+	-5	+						+	*	+		+	2
+	-6	+						+		+		+	(1)
Measr		+rater						* = 1		-round		Scale	

Fuente: Creada por la autora

Queda patente que los expertos consideraron que el contenido de los ítems tenía una distribución equilibrada a lo largo de la escala desde B1 del MCER hasta B2. Los umbrales Rasch-Andrich muestran que los expertos colocaron los ítems en distintas categorías con un aumento progresivo de dificultades, lo que concuerda con las pretensiones originales de los desarrolladores de la prueba. Los expertos discreparon ligeramente, y el experto 10 asignó los ítems a un nivel de dificultad levemente inferior al los demás expertos. No obstante, el análisis MFR toma este tipo de variaciones en cuenta y es capaz de dar una calificación a cada uno de los ítems mediante el empleo de un ‘promedio justo’, el cual realiza ajustes por la lenidad o severidad de los expertos. Si aplicamos los umbrales de Rasch Thurston, que miden el límite en el que un ítem tiene un 50 % de probabilidades de ser asignado a una categoría superior o inferior (i.e. el umbral de probabilidad acumulativa de 50 % (la mediana)), llegamos a unos puntos de corte iniciales de 13 para B1 y 34 para B2.

### **Método *Bookmark* de establecimiento de puntos de corte**

Antes de pasar al estudio principal, los participantes recibieron retroalimentación, que incluía datos de impacto, sobre el método *Basket*. Posteriormente, los expertos realizaron una discusión final, llegando a la conclusión general de que el punto de corte para B2 resultaba algo alto (solo 81 candidatos, o el 17%, habría conseguido el B2). En este punto, se les explicó que estos puntos de corte no serían los finales y que el ejercicio anterior se había realizado con el fin de permitir a los participantes que se familiarizaran con la prueba antes de ver los valores de dificultad reales que se iban a emplear para el establecimiento de los puntos de corte.

A continuación, los participantes recibieron una explicación del método *Bookmark*, además de una copia del CIO. Al terminar la primera ronda, las estimaciones de los expertos fueron presentadas al grupo, junto con los datos de impacto. Los expertos procedieron a discutir los resultados más a fondo antes de pasar a la segunda ronda, en la cual se les permitió cambiar la asignación del marcador. La tabla III muestra los resultados de las dos rondas<sup>1</sup>.

TABLA III. Resultados de los puntos de corte establecidos mediante el método *Bookmark*

I <sup>a</sup> Ronda	Número de páginas (número de expertos)	Media (DE)	Mediana	Habilidad ( $\theta$ ) después de ajustar a la PR de .67	Punto de corte final
B1	8/9 (3) 10/11 (2) 12/13 (2) 13/14 (1) 15/16 (1)	11 (2.87)	10/11	-.13 logits	22
B2	25/26 (1) 28/29 (2) 29/30 (5) 30/31 (1)	28.44 (1.42)	29/30	1.04 logits	29
<b>2<sup>a</sup> Ronda</b>					
B1	8/9 (7) 9/10 (1) 10/11 (1)	8.33 (.71)	8/9	-.70 logits	16
B2	28/29 (6) 29/30 (3)	28.56 (.73)	28/29	.93 logits	28

Fuente: Creada por la autora

La decisión global para el conjunto de los expertos se obtiene de la mediana del segundo grupo de marcadores de todos los participantes; para establecer el punto de corte, se recomienda emplear el valor zeta inferior (Consejo de Europa, 2009). La mediana en este caso corresponde a un nivel de capacidad de  $\theta = -.7$  logits para B1 y  $\theta = 0.93$  logits para B2, después de haber realizado ajustes por la PR de 0.67. Si aplicamos este nivel de capacidad a la curva de características de la prueba, resulta equivalente a una puntuación bruta de 16 para el B1 y de 28 para el B2.

Finalmente, los expertos realizaron una última discusión en la cual consideraron tanto los datos de impacto como aquellos de validez externa. Los participantes se mostraron satisfechos con las decisiones finales y de acuerdo con que los puntos de corte establecidos deberían ser utilizados para la prueba. Los subsiguientes datos de impacto para la convocatoria de marzo 2017 dan los siguientes resultados: 135 No aptos (el 29 %), 166 Aptos de nivel B1 (el 36 %) y 163 Aptos de nivel B2 (el 35 %).

## Validez del procedimiento

Los resultados del cuestionario posterior al proceso de establecimiento de puntos de corte se muestran en la tabla IV. Para cada una de las preguntas, los participantes contestaron usando una escala *Likert* de cuatro puntos, donde 1 representa total acuerdo. En términos generales, se evidenció un alto nivel de acuerdo, los miembros del panel confiaban en el procedimiento y en sus decisiones finales. En particular, todos los miembros consideraban que las decisiones de establecimiento de los puntos de corte finales representaban los niveles B1 y B2 del MCER para la destreza de comprensión oral de manera justa y precisa.

TABLA IV. Resultados del cuestionario

	Mean	SD
Creo que comprendo bien los niveles del MCER después de las sesiones de familiarización.	1.3	.48
Me han ayudado las explicaciones del constructo de la prueba y de sus especificaciones.	1.5	.71
He comprendido los procedimientos del método <i>Basket</i> para el establecimiento de puntos de corte.	1.2	.42
Confiaba en mis decisiones a la hora de contestar a la pregunta, '¿En qué nivel del MCER debe estar para poder contestar a este ítem?'	1.6	.52
He entendido el concepto de 'candidato de mínima capacidad'	1	.00
El uso del método <i>Basket</i> para el establecimiento de puntos de corte me ha ayudado a comprender los contenidos de la prueba.	1.2	.42
He comprendido los procedimientos del método <i>Bookmark</i> para el establecimiento de puntos de corte.	1.33	.71
He comprendido el concepto de PR y el significado de la probabilidad del 67%	1.11	.33
Confiaba en la colocación de mi marcador en la primera ronda.	1.67	.50
Confiaba en la colocación de mi marcador en la segunda ronda.	1.11	.33
Los puntos de corte finalmente recomendados por el grupo representan los niveles B1 y B2 del MCER para la destreza de comprensión oral de manera justa.	1	.00

Fuente: Creada por la autora

## Validez interna

La validez interna de los procedimientos para el establecimiento de los puntos de corte trata de la exactitud y la consistencia de los resultados así como de la calidad de las estimaciones determinadas y publicadas. Para el método *Basket*, los resultados mostraron que los expertos tenían una consistencia interna con *Infit* MNSQ medio de 0.98 y el DE de 0.28. Los *Infit* MNSQ y *Zstd* muestran la existencia de un experto con *overfit* y otro con *underfit*. No obstante, solo uno de los expertos muestra un ligero *underfit* en términos de *outfit* y había una alta similitud entre los resultados esperados y observados, como sería de esperar de evaluadores actuando como expertos independientes (Linacre, 2017).

Con el fin de reforzar la vinculación de la prueba con el MCER, las decisiones de los expertos sobre la dificultad de los ítems tomadas con el método *Basket* fueron comparadas con las dificultades observadas mediante el análisis *Rasch*. Este análisis muestra más evidencia en la calidad de la estimaciones de establecimiento de puntos de corte. El conjunto de datos es limitado, con rangos vinculados en los ítems estimados, y una correlación Tau de Kendall de  $\tau = .57, p < .001$  demuestra una relación significativa entre la estimación de la dificultad del ítem (empleando valores medianos de la segunda ronda) y los parámetros de dificultad *Rasch*. Este resultado proporciona evidencia de que los expertos reconocieron un aumento progresivo de la dificultad de los ítems.

Es necesario que proporcionemos evidencia de la exactitud de la clasificación de los puntos de corte. Los puntos de corte finalmente establecidos después de la segunda ronda del método *Bookmark* mostraron un alto grado de consenso, con DEs muy pequeños. La validez de estos puntos de corte puede ser evidenciada mediante el informe del error estándar de estimación ( $EE_E$ ), una estimación de las probabilidades de replicar los puntos de corte recomendados (Tannenbaum y Cho, 2014, p.245). Cohen, Kane y Crooks (1999, p.364) sostienen que el  $EE_E$  debería ser  $\leq 1/2 EEP$ . La tabla V muestra los resultados de este cálculo para las dos rondas.

TABLE V. Resultados para la exactitud de la clasificación

	Ronda 1		Ronda 2	
	B1	B2	B1	B2
Desviación estándar para el punto de corte medio (DE)	2.87	1.42	0.71	0.73
Error estándar de la prueba (EEP)	2.5	2.5	2.5	2.5
Error estándar de estimación ( $EE_E$ )	1.01	0.50	0.25	0.26
$\frac{EE_E}{EEP}$	0.4	0.2	0.1	0.1

Fuente: Creada por la autora

Se puede ver que la  $EE_E$  era siempre inferior a la mitad del EEP, y por tanto los puntos de corte cumplen el criterio de calidad. Sin embargo, es importante destacar aquí que se ha aducido que las estimaciones de la fase final del proceso no suelen exhibir mucha variabilidad debido a que esta parte del proceso anima a los expertos a converger y llegar a una decisión consensuada (Linn, 2003).

## Validez externa

La validez externa se refiere al uso y a la comparación de diferentes métodos para el establecimiento de puntos de corte, además de las comparaciones o triangulaciones con medidas obtenidas de otros estudios. Al igual que han informado previamente otros estudios, los dos procedimientos produjeron puntos de corte diferentes tanto para nivel B1 como para nivel B2. Para nivel B2, los resultados de la relación entre la puntuación bruta y los valores *zeta* muestran que existe cierto grado de discrepancia entre los dos métodos, una diferencia de poco más de *2 errores estándar*. A nivel B1, sin embargo, el punto de corte del método *Basket* está a menos de *1 error estándar* del punto de corte establecido con el método *Bookmark*, lo cual evidencia que los dos métodos produjeron resultados bastante consistentes para esta parte de la prueba, a pesar de que el método *Basket* solo se empleara para que los expertos se familiarizaran con el contenido.

Para aquellos candidatos que habían realizado cursos de familiarización del examen en el CLM antes de la convocatoria, se procedió a analizar sus

resultados con más detalle, mediante la comparación de sus notas finales con los resultados pronosticados por sus profesores. A pesar del número reducido de candidatos en cuestión, la comprobación de la exactitud de clasificación de estos candidatos permitió cierto grado de triangulación y aportó otra contribución más al conjunto de evidencias de validez. La tabla VI muestra los resultados.

Este estudio, de carácter limitado, y enfocado en el criterio de estimaciones de profesores, muestra que existe una alta correlación entre los pronósticos de los profesores y la puntuación obtenida en la prueba  $\tau = .85$ ,  $p < .001$ . El acuerdo de clasificación para todos los pronósticos de resultados B1 y B2 es del 100 %, lo cual nos proporciona algo más de evidencia de validez externa sobre la exactitud de los puntos de corte. Vemos que, de los candidatos a los que sus profesores pronosticaron un resultado en umbral entre los niveles de A2/B1 y B1/B2, alrededor del 50 % obtuvo el nivel superior. No obstante, cabe entender que el presente estudio es de una escala reducida, y que lo ideal sería otro estudio de mayor escala que pudiera proporcionar mejor evidencia de validez externa.

**TABLA VI.** Correspondencia entre los pronósticos de profesores y las decisiones finales basadas en los puntos de corte establecidos.

Nivel MCER pronosticado para la prueba de comprensión oral	Puntuación en la prueba	Nivel MCER recibido en la prueba
No Apto	1	No Apto
No Apto	11	No Apto
A2/B1	12	No Apto
A2/B1	12	No Apto
A2/B1	14	No Apto
A2/B1	17	B1
A2/B1	19	B1
B1	17	B1
B1	19	B1
B1	19	B1
B1	22	B1
B1	22	B1
B1	23	B1

B1	25	B1
B1	26	B1
B1	27	B1
B1	27	B1
B1/B2	26	B1
B1/B2	27	B1
B1/B2	33	B2
B1/B2	33	B2
B1/B2	36	B2
B2	30	B2
B2	31	B2
B2	31	B2
B2	35	B2
B2	36	B2
C1	40	B2

Fuente: Creada por la autora

## Análisis de los resultados

Este estudio ha proporcionado evidencia que fortalece afirmaciones sobre el uso e interpretación de los resultados de la prueba UGR B1/B2 en cuanto a la vinculación con el MCER se refiere. El proceso para el establecimiento de puntos de corte se ha detallado como parte elemental del proceso de vinculación y, de esta manera, queda documentado el seguimiento de un proceso razonable y sistemático para llegar a establecer unos puntos de corte recomendados. Se ha dado pruebas sólidas para justificar los puntos de corte, cuyo establecimiento se puede considerar una mezcla de estimaciones, sicometría, y pragmática (Hambleton y Pitonoak, 2006, p.435), y donde la cuestión no es que el punto de corte sea correcto o no, sino que las decisiones basadas en los puntos de corte sean razonables, aceptables en términos generales, y tengan consecuencias positivas (Kane, 2017, p.11).

El punto de partida del manual de adecuación con el MCER (Consejo de Europa, 2009) es, sin duda, el uso de estimaciones de expertos, una metodología muy recomendada y aplicada en investigaciones sobre el establecimiento de puntos de corte. No obstante, algunos estudios

han informado sobre las metodologías que hacen uso de estimaciones de expertos como no fiables, indudablemente como consecuencia del hecho de que el MCER no proporciona suficientes descripciones con precisión de sus niveles de capacidad (Alderson et ál., 2006; Fulcher, 2004; Weir, 2005). Existe la posibilidad de que los expertos en un estudio de adecuación interpreten los descriptores del MCER de manera diferente o que tengan interiorizada su propia idea de exactamente qué significa estar en una franja específica del MCER (Eckes, 2012; Harsch y Hartig, 2015; Papageorgiou, 2010). En efecto, North y Jones (2009) sostienen que ni la familiarización con el MCER, ni la estandarización, ni las estimaciones de índices de consistencia y acuerdo, podrán demostrar que un grupo particular de expertos involucrados en la estimación de niveles para un determinado idioma no lleven consigo a la tarea sus propias interpretaciones, fruto de su particular perspectiva cultural (p.16).

No obstante, y a pesar de estas limitaciones, los expertos en el presente estudio se mostraron satisfechos con sus resultados y consideran que los puntos de corte son representativos de un candidato que pueda realizar tareas asociadas con el nivel. Por tanto, el estudio ha proporcionado más evidencia hacia la validez sustantiva y la validez de constructo de la prueba. Las dos metodologías empleadas requirieron discusiones extensas sobre la relación de la dificultad de ítems individuales con el MCER; como tal, refuerzan la afirmación de que el contenido de la prueba se adecúa con el MCER, en la línea de Kanistra y Harsch (2017) en su estudio de vinculación sobre la prueba ISE de Trinity. De hecho, el problema de cómo definir al candidato que actúe en un nivel mínimo de actuación resulta ser una tarea mucho más fácil para una prueba que ha sido desarrollada para poner en práctica el modelo del MCER; debido a que el estándar ya ha sido incorporado en la prueba, los puntos de corte adquieren un carácter más significativo (Tannenbaum y Wylie, 2008). En cambio, un estudio de vinculación retroactiva tendría que prestar particular atención a la fase de especificación de contenidos del proceso de adecuación con el MCER para que proporcione evidencia de que los contenidos de la prueba miden las capacidades lingüísticas descritas por el marco (Tannenbaum y Cho, 2014). Varios estudios de vinculación al MCER han señalado problemas a la hora de relacionar contenidos con los descriptores (p.ej. Brunfaut y Harding, 2014). Además, se ha argumentado que son muy pocos los creadores de pruebas que prestan atención a esta fase del proceso de vinculación con el MCER (Green, 2017); si los

contenidos no se alinean con el MCER, pues existe poca justificación para la realización de un estudio de establecimiento de puntos de corte. Aquí cabe destacar que la prueba de la UGR ha sido desarrollada de forma específica para que sea representativa del MCER; como tal, el estudio de vinculación no tiene carácter retroactivo, sino que forma parte del mismo proyecto de desarrollo de la prueba, y el estándar está ya incorporado a la prueba de forma implícita. Las especificaciones de la prueba incluyen descriptores del MCER y las directrices para los desarrolladores de ítems abarcan varias categorías presentadas en el Consejo de Europa (2009). Dicho proceso tiene parecido con el enfoque *a priori* de la prueba *Pearson Test of English Academic* (De Jong y Zheng, 2016).

En cuanto a la metodología empleada para el estudio, no existe el mejor método para el establecimiento de puntos de corte; los métodos deben adecuarse a la situación. Como ya se ha observado, el estudio principal, el del método *Bookmark*, se considera apropiado para usar en contextos de evaluación que se basan en el modelo *Rasch* para el desarrollo de pruebas. Por otra parte, el empleo del método *Basket* también tuvo un valor inestimable durante la primera fase del estudio, ya que permitió a los expertos que se familiarizaran con los contenidos de la prueba. En este caso, habría sido imposible comenzar con el estudio *Bookmark* de inmediato, y esto resulta de particular relevancia para una prueba de comprensión oral. El CIO no presenta los ítems en el mismo orden que la prueba original; por lo tanto, pasar por los archivos de audio de manera artificial para encontrar y tratar cada uno de los ítems individuales, sin haber tenido conocimientos previos de la prueba, habría sido de suma dificultad. El Consejo de Europa (2009) recomienda que, para una prueba de comprensión oral, los expertos deben estar provistos de un ordenador para poder hacer esta tarea. En cambio, es la firme convicción de la autora de que la metodología empleada en este estudio supone una solución mucho más pragmática al problema, y que cualquier estudio que emplee el método *Bookmark* debe realizar un estudio parecido al del método *Basket* primero. Un enfoque similar fue adoptado por Harsch y Hartig (2015), aunque su razonamiento era distinto; querían desvincular la tarea de separación de contenidos y la tarea de emparejamiento de dificultad de los ítems porque las consideraban tareas que requerían tipos de estimación muy diferentes.

De hecho, un resumen de estudios en los que se empleó el método *Bookmark* (Peterson, Schulz, y Engelhard, 2011) concluyó que los

expertos confiaban en los puntos de corte resultantes. Indudablemente, los puntos de corte del presente estudio cuentan con el firme respaldo de la evidencia de validez externa que ha sido proporcionada. Dicha evidencia necesitará fortalecerse con su continuación en otros estudios en el futuro; con el tiempo, se podrían recopilar muestreos más amplios y quizás también proceder al empleo de un método del tipo *Prototype group*, similar al del estudio realizado por Eckes (2012).

Otro comentario que cabe mencionar es el hecho de que algunos ítems del CIO tenían unos valores empíricos de dificultad que no coincidían con las estimaciones de los expertos. Este fenómeno también se ha encontrado en estudios previos (p.ej. Figueras, Kaftandjieva, y Takala, 2013). Además, se ha señalado ampliamente que a menudo los expertos no son capaces de identificar la dificultad de un ítem de forma correcta (Alderson, 1993). El presente estudio no está libre de este problema, como evidencia la divergencia de los puntos de corte para el nivel B2 obtenidos por los métodos *Basket* y *Bookmark*. No obstante, es la convicción de la autora que dicha discrepancia se debe al hecho de que los participantes ya habían recibido mucha más información cuando tomaron sus decisiones finales empleando el método *Bookmark*. Como consecuencia, ellos pudieron tomar decisiones más informadas que tuvieron en cuenta las dificultades reales de los ítems, y no solo estimaciones subjetivas basadas en su opinión. Para el punto de corte a nivel B1, los expertos colocaron el marcador en un punto algo bajo del continuo de dificultad, debido a la presencia de ítems que creían ser más representativos de capacidades en umbrales entre B1 y B2. No obstante, después de aplicar el ajuste por la PR, el punto de corte para el nivel B1 obtenido por los dos métodos de establecimiento era muy parecido (menos de un error estándar), lo cual es una confirmación más de que la recopilación de múltiples fuentes de evidencia puede aumentar el nivel de confianza que tenemos en la decisiones cualitativas que tomamos.

## Conclusión

Respecto al enfoque del argumento por la validez presentado por ALTE, el desarrollo de la prueba de la UGR se ha atendido a unas especificaciones detalladas y basadas en el MCER (inferencia de observación). Muestra *validez de puntuación* hasta el punto de que todas las tareas han sido

pilotadas y que los ítems muestran buenas propiedades sicométricas (inferencia de evaluación). El presente estudio ha reforzado la inferencia de extrapolación; se puede considerar que los puntos de corte son traducibles y consistentes y que los resultados se pueden emplear para establecer el nivel de capacidad en todas las futuras versiones de la prueba. Mediante el uso de un sistema de equiparación de ítems-ancla comunes (véase Kolen y Brennan, 2014; North y Jones, 2009; Wright y Stone, 1979), todas las tareas de la prueba de la UGR se pueden calibrar según la misma escala de medición *Rasch*; por consiguiente, la dificultad de la prueba será la misma en cada una de sus versiones, lo que respalda la inferencia de generalización.

Los puntos de corte no solo tienen consecuencias para los candidatos sino también para un amplio grupo de otras partes interesadas, como son los padres, educadores y legisladores. El establecimiento de puntos de corte es una parte fundamental del proceso de desarrollo de pruebas; por consiguiente, no solo debe tratarse como si fuera una actividad aislada sino también como un componente esencial e integral del continuado proceso de validación (Papageorgiou y Tannenbaum, 2016). En el contexto de la prueba de la UGR, esto incluye la implantación de futuros estudios para controlar las consecuencias de la aplicación de los puntos de corte que han sido decididos en el presente estudio, algo que forma parte de todo un abanico de controles de validez implantados por los desarrolladores de la prueba. De igual manera, se ha realizado un proceso de establecimiento de puntos de corte para la prueba de comprensión escrita de la UGR, y otros estudios parecidos se realizan con regularidad como parte de las propias sesiones internas de *benchmarking* (homogeneización de criterios) y formación en el CLM para las pruebas de producción escrita y oral. En efecto, la validación de la interpretación y uso de los resultados de la prueba binivel B1/B2 de la UGR es un proceso continuo: es precisamente esta recopilación de evidencia de validez la que proporciona el necesario respaldo para que todas las partes interesadas confíen en cualquier decisión que se fundamente en los resultados.

En España, los exámenes *CertAcles* han contribuido hasta cierto punto a responder a las recomendaciones hechas por Halbach, Lafuente y Guerra (2013) respecto a la acreditación de lenguas y la homogeneización de criterios. Estas pruebas están diseñadas para que sean consistentes con el MCER, y exigen la provisión de evidencia de fiabilidad y validez.

No obstante, como advierten Deygers et ál. (2017), la presión sobre los organismos de evaluación para que demuestren que sus pruebas se han adecuado con el MCER ha conducido al mal uso de pruebas en muchos contextos donde dicha afirmación no está corroborada. A tal respecto, sugeriría a todos los creadores y administradores de pruebas en el contexto actual que lleven a cabo estudios de establecimiento de puntos de corte para que los resultados oficiales de sus pruebas sean más significativos. Aunque muchas instituciones han visto el punto de corte como un estándar normativo (p.ej. el 60 %), los desarrolladores de pruebas deben cambiar esa perspectiva hacia el uso de puntos de corte que se puedan defender como representación del estándar de dominio, para que así podamos fortalecer la afirmación de que nuestras pruebas gozan de una buena vinculación con el MCER.

## Referencias bibliográficas

- AERA (American Educational Research Association), APA (American Psychological Association) & NCME (National Council on Measurement in Education). (2014). *Standards for educational and psychological testing*. Washington, DC: AERA.
- Alderson, J. C. (1993). Judgements in language testing. In D. Douglas & C. Chapelle (Eds.), *A new decade of language testing research* (pp. 46-57). Alexandria, VA: TESOL.
- (2012). Principles and Practice in Language Testing: Compliance or Conflict? Presentation at TEA SIG Conference: Innsbruck. Recuperado de <http://tea.iatefl.org/inns.html>
- Alderson, J.C., Figueras, N, Kuijper, H., Nold, G, Takala, S & Tardieu, C. (2006). Analysing Tests of Reading and Listening in Relation to the Common European Framework of Reference: The Experience of The Dutch CEFR Construct Project, *Language Assessment Quarterly*, 3(1), 3-30.
- ALTE/Council of Europe (2011) Manual for Language Test Development and Examining. For use with the CEFR. Recuperado de [http://www.coe.int/t/dg4/linguistic/ManualLangageTest-Alte2011\\_EN.pdf](http://www.coe.int/t/dg4/linguistic/ManualLangageTest-Alte2011_EN.pdf)

- Bachman, L. F. (2005). Building and supporting a case for test use. *Language Assessment Quarterly* 2(1), 1-34.
- Bachman, L.F., & Palmer, A. (2010). *Language assessment in practice*. Oxford: Oxford University Press.
- Brunfaut, T., & Harding, L. (2014). *Linking the GEPT listening test to the Common European Framework of Reference*. Taiwan: Language Training and Testing Centre.
- Chapelle, C. A. (2012). Validity argument for language assessment: The framework is simple... *Language Testing*, 29(1), 19-27.
- Cizek, G. J. (2012). The forms and functions of evaluations in the standard setting process. In G. J. Cizek (Ed.), *Setting performance standards: Foundations, methods, and innovations* (2<sup>nd</sup> ed., pp. 165 - 178). New York: Routledge.
- (2016). Validating test score meaning and defending test score use: Different aims, different methods. *Assessment in Education: Principles, Policy & Practice*, 23(2), 212-225.
- Cohen, A. S., Kane, M. T., & Crooks, T. J. (1999). A generalized examinee-centered method for setting standards on achievement tests. *Applied Measurement in Education*, 12(4), 343-366.
- Conferencia De Rectores De Las Universidades Españolas (CRUE). (2011, 8 de septiembre). Propuestas sobre la acreditación de idiomas. Recuperado de <http://www.acreditacion.crue.org/>
- Council of Europe. (2008). *Recommendation CM/Rec (2008)7 of the Committee of Ministers to member states on the use of the Council of Europe's Common European Framework of Reference for Languages (CEFR) and the promotion of plurilingualism*. Strasbourg, France: Council of Europe. Recuperado de [http://www.coe.int/t/dg4/linguistic/Conventions\\_EN.asp](http://www.coe.int/t/dg4/linguistic/Conventions_EN.asp)
- (2009). Relating language examinations to the Common European Framework of Reference for Languages: Learning, teaching, assessment. Strasbourg, France: Council of Europe. Recuperado de [http://www.coe.int/T/DG4/Linguistic/Manuel1\\_EN.asp](http://www.coe.int/T/DG4/Linguistic/Manuel1_EN.asp)
- De Jong, J. & Zheng, Y. (2016) Linking to the CEFR: validation using a priori and a posteriori evidence. In, Banerjee, J. and Tsagari, D. (eds.) *Contemporary Second Language Assessment*. London, GB, Bloomsbury Academic pp. 83-100. (Contemporary Applied Linguistics, 4).

- Deygers, B., Zeidler, B., Vilcu, D., & Carlsen, C. H. (2017). One Framework to Unite Them All? Use of the CEFR in European University Entrance Policies. *Language Assessment Quarterly*, 1-13.
- Downey, N. & Kollias, C. (2010). Mapping the Advanced Level Certificate in English (ALCE™) examination onto the CEFR. Aligning Tests with the CEFR, Reflections on using the Council of Europe's draft Manual, Martyniuk, W. (ed). Cambridge University Press. Cambridge. 119-129.
- Eckes, T. (2009). Many-facet Rasch measurement. In S. Takala (Ed.), *Reference supplement to the manual for relating language examinations to the Common European Framework of Reference for Languages: Learning, teaching, assessment* (Section H). Strasbourg, France: Council of Europe/Language Policy Division.
- (2012). Examinee-centered standard setting for large-scale assessments: The prototype group method. *Psychological Test and Assessment Modeling*, 54, 257-283.
- Ferrara, S., & Lewis, D. (2012). The Item-Descriptor (ID) Matching method. In G. J. Cizek (Ed.), *Setting performance standards: Foundations, methods, and innovations* (2nd ed., pp. 255-282). New York: Routledge.
- Figueras, N. (2012). The impact of the CEFR. *ELT journal*, 66(4), 477-485.
- Figueras, N., Kaftandjieva, F., & Takala, S. (2013). Relating a Reading Comprehension Test to the CEFR Levels: A Case of Standard Setting in Practice with Focus on Judges and Items. *Canadian Modern Language Review/La Revue Canadienne Des Langues Vivantes*, 69(4), 359-385.
- Fulcher, G. (2004). "Deluded by artifices? The Common European Framework and harmonization." *Language Assessment Quarterly*, 1(4), 253 - 266.
- Fulcher, G., & Owen, N. (2016). Dealing with the demands of language testing and assessment. *The Routledge Handbook of English Language Teaching*, 109-120.
- Green, A. (2017). Linking Tests of English for Academic Purposes to the CEFR: The Score User's Perspective, *Language Assessment Quarterly*, 1-16.
- Halbach, A., Lazaro Lafuente, A., & Perez Guerra, J. (2013). La lengua inglesa en la nueva universidad española del EEES: The role of the English language in post-Bologna Spanish universities. *Revista de Educación*, 362, 105-132.
- Hambleton, R., & Jirka, S. (2006) Anchor-based methods for judgmentally estimating item statistics. In S. Downing and T. Haladyna (Eds.), *Handbook of test development* (pp. 399-420). Mahwah, Nj: Erlbaum.

- Hambleton, R., & Pitoniak, M. J. (2006). Setting performance standards. In R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 433–470). Westport, CT: American Council on Education/ Praeger Publishers.
- Harsch, C. & Hartig, J. (2015). What Are We Aligning Tests to When We Report Test Alignment to the CEFR?, *Language Assessment Quarterly*, 12:4, 333-362.
- Kaftandijeva, F. (2010). *Methods for setting cut scores in criterion-referenced achievement tests. A comparative analysis of six recent methods with an application to tests of reading in EFL*. Cito, Arnhem: EALTA. Recuperado de [www.ealta.eu.org/documentsresources/FK\\_second\\_doctorate.pdf](http://www.ealta.eu.org/documentsresources/FK_second_doctorate.pdf)
- Kane, M. (2012). Validating score interpretations and uses. *Language Testing*, 29 (1), 3 –17.
- (2017). Using Empirical Results to Validate Performance Standards BT - Standard Setting in Education: The Nordic Countries in an International Perspective. In S. Blömeke & J.-E. Gustafsson (Eds.), (pp. 11–29). Cham: Springer International Publishing.
- Kanistra, V. & Harsch, C. (2017). *Using the Item Descriptor Matching method to enhance validity when aligning test to the CEFR*. 4th Meeting of the EALTA CEFR Special Interest Group. Recuperado de [http://www.ealta.eu.org/events/SIG\\_CEFR\\_london2017/presentations/Presentations/1400-1530/CEFR%20SIG%20Voula%20and%20Claudia.pdf](http://www.ealta.eu.org/events/SIG_CEFR_london2017/presentations/Presentations/1400-1530/CEFR%20SIG%20Voula%20and%20Claudia.pdf)
- Kantarcioğlu, E, Thomos, C, O'Dwyer, J & O'Sullivan, B. (2010) 'Benchmarking a high-stakes proficiency exam: the COPE linking project', in W. Martyniuk (ed) *Aligning Tests with the CEFR: reflections on using the Council of Europe's draft Manual*. Cambridge University Press. Cambridge.
- Kolen, M. J., & Brennan, R. L. (2014). Test equating, scaling, and linking: methods and practices (3rd ed.). New York: Springer-Verlag.
- Linacre, J. M. (2017). Facets computer program for many-facet Rasch measurement, version 3.80.0. Beaverton, Oregon: Winsteps.com
- Linn, R. L. (2003). Performance Standards: Utility for Different Uses of Assessments. *Education policy analysis archives*, 11, 31.
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13–103). New York: Macmillan.
- Mitzel, H.C., Lewis, D.M., Patz, R.J., & Green, D.R. (2001). The bookmark procedure: Psychological perspectives. In G.J. Cizek (Ed), *Setting*

- performance standards: Concepts, methods and perspectives (pp. 249-281). Mahwah, NJ: Lawrence Erlbaum Assoc.
- North, B., & Jones, N. (2009). Relating language examinations to the Common European Framework of Reference for Languages. Further material on maintaining standards across languages, contexts and administrations by exploiting teacher judgment and IRT scaling. *Strasbourg: Language Policy Division*.
- Papageorgiou, S. (2010). Investigating the decision-making process of standard setting participants. *Language Testing*, 27(2), 261–282.
- Papageorgiou, S., & Tannenbaum, R. J. (2016). Situating Standard Setting within Argument-Based Validity. *Language Assessment Quarterly*, 13(2), 109–123.
- Peterson, C. H., Schulz, E. M. and Engelhard Jr., G. (2011), Reliability and Validity of Bookmark-Based Methods for Standard Setting: Comparisons to Angoff-Based Methods in the National Assessment of Educational Progress. *Educational Measurement: Issues and Practice*, 30: 3–14.
- Reckase, M. D. (2010). NCME 2009 presidential address: What I think I know. *Educational Measurement: Issues and Practice*, 29(3), 3–7.
- Tannenbaum, R. J. & Cho, Y. (2014). Critical Factors to Consider in Evaluating Standard-Setting Studies to Map Language Test Scores to Frameworks of Language Proficiency, *Language Assessment Quarterly*, 11(3), 233-249.
- Tannenbaum, R. J. & Wiley E. C. (2008). Linking English-language test scores onto the Common European Framework of Reference: An application of standard-setting methodology. *ETS Research Report Series*, 2008(1).
- Weir, C. J. (2005). Limitations of the council of Europe's framework of reference (CEFR) in developing comparable examinations and tests. *Language Testing*, 22(3), 281–300.
- Wright, B. D., & Stone, M. H. (1979). *Best test design*. Chicago: MESA.

**Información de contacto:** Caroline Shackleton. Centro de Lenguas Modernas de la Universidad de Granada. Placeta del Hospicio Viejo s/n 18009, Granada, Spain. E-mail: csarah@ugr.es



# Incidencia de la crisis económica en la segregación escolar en España<sup>1</sup>

## Impact of the economic crisis on school segregation in Spain

DOI: 10.4438/1988-592X-RE-2017-381-381

F. Javier Murillo  
Cynthia Martínez-Garrido  
*Universidad Autónoma de Madrid*

### Resumen

¿La grave crisis económica que afecta a España desde 2008 ha tenido su repercusión en la equidad del sistema educativo, concretamente en la magnitud de la segregación escolar? Este estudio tiene por objetivo determinar la evolución en los últimos 15 años de la segregación escolar de carácter socioeconómico y por origen nacional en España y las Comunidades Autónomas. Para ello, se realiza una explotación especial de los datos disponibles del estudio PISA desde el año 2000 hasta 2015 y se estiman cuatro índices de segregación: índice de Gorard, de Disimilitud, de Hutchens y de Aislamiento. La segregación socioeconómica se calcula para el 25% de estudiantes con familias de menor socioeconómico y cultural (Q1) y el 25% mayor (Q4), y la segregación por origen nacional se estima considerando como grupo minoritario a los estudiantes nacidos en el extranjero. Los resultados indican que la segregación escolar por nivel socioeconómico ha disminuido ligeramente desde 2000 hasta 2012, y ha aumentado fuertemente hasta 2015 (un 3,1% para Q1 y 3,6% Q4). La segregación escolar por origen nacional, acompañada por el aumento del porcentaje de extranjeros, ha disminuido desde el 2000 de forma importante hasta 2009. Entonces se ha producido un estancamiento a pesar del crecimiento del porcentaje de inmigrantes. De esta

---

<sup>(1)</sup> Este artículo ha sido desarrollado en el marco del Proyecto I+D+i de Excelencia "Escuelas en contextos socio-económicamente desafiantes: Una aproximación desde la Educación para la Justicia Social de Investigación" Ref: EDU2014-56118-P. Financiado con fondos del Ministerio de Economía y Competitividad-Gobierno de España.

forma, se puede afirmar que hay evidencias de que la crisis ha incidido en el aumento de la segregación escolar en España, rompiendo la tendencia a la disminución desde 2000. Urge, por tanto, que las Administraciones educativas tomen medidas que reviertan esa situación y que encaminen al sistema educativo en esa tendencia a la baja que había antes de la crisis. Las diferencias encontradas por Comunidades indican que la política educativa pública puede profundizar en la construcción de un sistema educativo equitativo o, por el contrario, que contribuya a profundizar las desigualdades sociales.

*Palabras clave:* Escuela, Evolución, Equidad, Recesión económica, Inmigración, Condiciones económicas, Integración escolar.

### **Abstract**

The equity of the education system, in particular, the magnitude of school segregation has been affected by the severe economic crisis in Spain since 2008? The aim of this study is to determine the evolution across the last 15 years of the socioeconomic school segregation and immigrant school segregation in Spain and its Autonomous Communities. To achieve our aim, we conduct a special exploitation of the available PISA dataset from 2000 to 2015; we estimated four segregation indexes: Gorard index, Dissimilarity index, Hutchens index, and Isolation index. The socioeconomic school segregation considerer as a minority group the 25% of students from families with lower socioeconomic and cultural level (Q1), as well as the 25% with higher socioeconomic and cultural level (Q4). The minority group considered to estimate the immigrant school segregation was the students born abroad. According to our data, the socioeconomic school segregation in Spain has slightly decreased from 2000 to 2012, and increased strongly to 2015 (3.1% for students from Q1, and 3.6% students from Q4). Along the increases of the percentage of immigrant students from 2000 to 2009, the immigrant school segregation has decreased. From this moment, the immigrant school segregation has remained, even though the percentage of immigrant students keep growing. These results, let us confirm that there is evidence of the impact that economic crisis has into the increase of the school segregation in Spain. This impact stemming the decrease trend since 2000. A greater effort must be made by the educational administrations to terminate this state of affairs. They should strive to pursue the decreased trend that already existed before economic crisis came. The differences founded between the Autonomous Communities let us know that the public education policy can help to develop an equitable educational system, or it can serve to heighten social inequalities.

*Keywords:* School, Evolution, Equity, Economic recession, Immigration, Economic situation, School integration.

## Introducción

En el año 2008 comenzó en España una crisis económica que afectó, y sigue afectando, a la vida de millones de personas. No es muy arriesgado afirmar que el sistema educativo es una de las instituciones sociales sobre las que más han repercutido la recesión y las decisiones que amparándose en ella se han tomado. Así, las importantes restricciones presupuestarias en este ámbito –de un 13% de la inversión pública en educación (Mas Ivars, Pérez García y Uriel, 2015)– han tenido serias consecuencias. Por citar algunos ejemplos, desde el año 2009 y a pesar del aumento de la matrícula, hay entre 28.000 y 32.000 docentes menos, han empeorado las condiciones económicas y laborales de los y las docentes, o se ha incrementado la ratio de alumnos por aula, afectando cinco veces más a la escuela pública que al sector educativo privado. Pero la crisis también ha tenido su impacto en las familias, especialmente en los hogares con niños y niñas que han visto mermados sus ingresos de forma importante; tanto, que se estima que en la actualidad uno de cada tres niños en España está en riesgo de pobreza y exclusión (Ayllón, 2015), lo que conlleva fuertes implicaciones en el sistema educativo (García-Gómez y Cabanillas, 2017).

No cabe duda que una condición necesaria, aunque sin duda no suficiente, para conseguir una sociedad más cohesionada y justa es tener un sistema educativo con las mismas características. La segregación escolar es un factor de inequidad educativa que atenta contra la igualdad de oportunidades e incide de forma directa en la generación de inequidades sociales. ¿La crisis ha afectado a la segregación escolar en España? No lo sabemos. Tenemos algunas evidencias de la segregación escolar en España en momentos puntuales (p. ej., Ferrer, 2008; Mancebón-Torrubia y Pérez-Ximénez, 2009; Murillo y Martínez-Garrido, 2018), pero hasta ahora no se ha realizado un estudio evolutivo que permita analizar lo ocurrido estos últimos años.

Así, este estudio tiene por objetivo determinar la evolución de la segregación escolar de carácter socioeconómico y por origen nacional en España y las Comunidades Autónomas utilizando los datos de la evaluación internacional PISA, para determinar la incidencia de la crisis económica.

## Revisión de la literatura

Se entiende segregación escolar como la distribución desigual de los estudiantes en las escuelas en función de sus características personales, sociales o de su condición. Con ello, como afirman Benito y González-Baetbó (2007, p. 41), se genera una “situación de homogeneidad social intraescolar y de heterogeneidad social interescolar”.

Las primeras investigaciones que estimaron la segregación escolar tuvieron como objetivo verificar el cumplimiento de la sentencia del llamado *caso Brown* por el cual, el 17 de mayo de 1954, la Corte Suprema de los Estados Unidos declaró inconstitucional la segregación por razón de color de piel en las escuelas públicas norteamericanas. Así, en la segunda mitad de los años 50 comenzó una fructífera línea de investigación que buscaba estimar la magnitud de la segregación étnica en las escuelas (Orfield, Ee, Frankenberg y Siegel-Hawley, 2016; Reardon y Owens, 2014; Stroub y Richards, 2013). Esta preocupación por la desigual distribución de los estudiantes se trasladó al estudio de la segregación por origen nacional de los estudiantes (p.ej., Cebolla-Boado y Garrido Medina, 2011; Stephan, 2013; Van-Zanten, 2007), y en los años 80 al estudio de la segregación escolar de carácter socioeconómica (p. ej., Gorard, Hordosy y See, 2013; Owens, Reardon y Jencks, 2016)

En España, las investigaciones que estiman la magnitud de la segregación escolar son recientes, de hace apenas 10 años, y se centran tanto en la segregación escolar de carácter socioeconómico y cultural como en la segregación por origen nacional. La Tabla I presenta una panorámica general de los trabajos más importantes.

**TABLA I.** Principales investigaciones que estiman la magnitud de la segregación escolar en España

ESTUDIO	TIPO DE SEGREGACIÓN	POBLACIÓN	BASE DE DATOS	ÍNDICE
Benito y González-Baetbó (2007)	Estudios de los padres Origen nacional	Municipios en Cataluña	Generalitat de Catalunya 2004	G, P
Sánchez Hugalde (2007)	Origen nacional	Municipios en Cataluña	Generalitat de Catalunya 2002	D
Alegre, Benito y González-Baetbó (2008)	Origen nacional	Municipios en Cataluña	Generalitat de Catalunya 2005	G
Ferrer (2008)	Nivel socioeconómico y cultural Estudios de los padres	España y Cataluña	PISA 2003	D
Síndic (2008)	Origen nacional	Cataluña	Generalitat de Catalunya desde 2001 a 2006	D, A, H
Valiente (2008)	Origen nacional	Cataluña	Generalitat de Catalunya 2001 y 2006	D, A, H
Mancebón-Torrubia y Pérez-Ximénez (2009)	Estudios de los padres Origen nacional	España	PISA 2006	D, C
Mancebón-Torrubia y Pérez-Ximénez (2010)	Nivel socioeconómico y cultural	España	PISA 2006	D
Ferrer, Castejón, Castel y Zancajo (2011)	Nivel socioeconómico y cultural	España y CCAA	PISA 2009	D
Bonal (2012a, b)	Origen nacional	Municipios en Cataluña	Generalitat de Catalunya 2010	D
Murillo, Martínez-Garrido y Belavi (2017)	Origen nacional	España y CCAA	PISA 2015	G
Murillo y Martínez-Garrido (2018)	Nivel socioeconómico y cultural	España y CCAA	PISA 2015	G, A

Nota: A: índice de Aislamiento. C: índice de Contacto. D: índice de Disimilitud. G: índice de Gorard. H: Índice de Hutchens. P: índice de Polarización.

Fuente: Elaboración propia.

La investigación que ha buscado estimar la magnitud de la segregación escolar en España ha abordado el estudio de tres tipos de segregación. En primer lugar, un buen número de trabajos abordan la segregación escolar de carácter socioeconómico y cultural; es decir, la distribución desigual

de los estudiantes en las escuelas en función del nivel socioeconómico y cultural de las familias. Es el caso de los trabajos de Ferrer (2008), Ferrer y otros (2011), Mancebón-Torrubia y Pérez-Ximénez (2010) y Murillo y Martínez-Garrido (2018). Otras investigaciones estudian la diferente distribución de los niños y niñas en los centros según el nivel educativo de sus padres (Benito y González-Baetbó, 2007; Ferrer, 2008; Mancebón-Torrubia y Pérez-Ximénez, 2009). Un último grupo de trabajos, los más numerosos, se ocupan de la segregación escolar por origen nacional; esto es, determinar la posible concentración en algunas escuelas de los estudiantes que no han nacido en España (Alegre et ál., 2008; Benito y González-Baetbó, 2007; Bonal, 2012a, 2012b; Mancebón-Torrubia y Pérez-Ximénez, 2009; Murillo et ál., 2017; Sánchez Hugalde, 2007; Síndic, 2008; Valiente, 2008).

El estudio de la magnitud de la segregación exige contar con grandes muestras estadísticamente representativas. Quizá por esa razón, los trabajos se han centrado en estudiar la situación del total estatal y, en la medida de la disponibilidad de los datos, de las Comunidades utilizando las distintas ediciones del estudio PISA (2003, 2006, 2009 y 2015), o de Cataluña, a partir de los datos del *Departament d'Educació de la Generalitat de Catalunya*. Ojalá que la actual obsesión evaluadora también se traduzca en aportar datos a la comunidad científica para poder realizar trabajos que permitan conocer mejor la educación española.

No existe consenso entre los investigadores e investigadoras españoles, sin embargo, en torno a qué índice utilizar para estudiar la magnitud de la segregación escolar. El índice más popular es el índice de Disimilitud (Duncan y Duncan, 1955), que refleja la proporción de estudiantes de grupo minoritario que debería cambiar de escuela para que exista una distribución homogénea entre ellas. Pero también son usados el índice de Gorard (2009), una variación del índice de Disimilitud que controla el tamaño del grupo minoritario y que se interpreta como la tasa de estudiantes que tendrían que cambiar de escuela para lograr su distribución igualitaria dentro del área geográfica de análisis, o el índice de Hutchens (2004) que estima la suma del alejamiento de cada escuela de la igualdad distributiva, es decir, la distancia entre la media geométrica de las participaciones de estudiantes del grupo minoritario en ausencia de segregación y la media geométrica de las participaciones reales. Menos habitual es el uso del índice de Polarización propuesto por Benito y González-Baetbó (2007) que estima la distribución de un

subgrupo de alumnos entre el 20% de escuelas con la mayor y menor proporción de ellos en determinado municipio.

Estos cinco índices miden lo que se entiende como dimensión de uniformidad de la segregación: el grado de desbalance con que los miembros del grupo minoritario se distribuyen entre las escuelas. La otra dimensión de la segregación escolar es la de Exposición, esto es, la probabilidad de encontrarse en una escuela con alguien de su mismo grupo. Los trabajos en España usan el índice de Aislamiento (Lieberman, 1981) o, de forma más excepcional, el índice de Contacto (Coleman, Hoffer y Kilgore, 1982).

La dificultad que surge a raíz del desacuerdo en el uso de un índice u otro es que cada uno tiene diferente significado, y aunque están altamente relacionados entre sí (Murillo, 2016), algunos aportan estimaciones más altas que otros, por lo que no son comparables. Así, de los cuatro más usados, el de Disimilitud es el que aporta cifras más altas, después el de Gorard y por último el de Hutchens. En el caso del índice de Aislamiento, además de medir otra dimensión, sus resultados están muy marcados por el tamaño del grupo minoritario. A ello hay que añadir que para estimar la segregación escolar hay que definir el grupo minoritario a estudiar, por lo que según se tome uno u otro, las cifras varían.

Analizando los hallazgos de estas investigaciones, se encuentra que la segregación escolar por nivel socioeconómico en España según los datos de PISA 2003 es de 0,40 tomando como grupo minoritario al 25% de los estudiantes con menor nivel socioeconómico (Ferrer, 2008). Con datos PISA 2006, Mancebón-Torrubia y Pérez-Ximénez (2010) lo estiman en 0,39, en este caso los autores consideran como grupo mayoritario tan sólo al 25% de los estudiantes con mayores recursos –valores medidos a través del índice de Disimilitud. Más recientes son los resultados de Murillo y Martínez-Garrido (2018) utilizando PISA 2015, quienes estiman la segregación en 0,38 (índice de Gorard) y 0,32 (índice de Aislamiento) –los autores promedian la segregación del 25% de los estudiantes con mayor y menor nivel socioeconómico. En Cataluña, el análisis de Benito y González-Baetbó (2007) señala que la segregación en el año 2004 varía entre 0,20 y 0,66, en función del municipio (valores medidos a través del índice de Gorard).

La segregación escolar en función del nivel de estudios de los padres en España según el estudio de Ferrer (2008) con datos PISA 2003 es de 0,34, usando como grupo minoritario lo estudiantes cuyos padres no

tienen estudios o no concluyeron la primaria y como grupo mayoritario el resto. Mancebón-Torrubia y Pérez-Ximénez (2009), utilizando PISA 2006, encontraron una segregación de 0,29; en este caso, el grupo minoritario lo conforman aquellos estudiantes cuyos padres no tienen estudios o sólo estudios primarios, y como grupo mayoritario aquellos cuyos padres tienen estudios superiores a los básicos. En Cataluña la segregación puede variar de 0,45 para aquellos estudiantes cuyos padres no tienen estudios o tienen incompleta la educación obligatoria, a 0,37 si los padres tienen estudios universitarios según el estudio realizado por Benito y González-Baetbó (2007) –valores promedio entre los municipios participantes en el estudio. Valores medidos con el índice de Disimilitud.

Por último, la segregación escolar del alumnado inmigrante en los centros públicos en España según Mancebón-Torrubia y Pérez-Ximénez (2009), utilizando los datos PISA, en el año 2006 el valor de la segregación en los centros públicos es 0,53, de 0,57 si se considera el total de los centros (valores medidos con el índice de Disimilitud). Con datos PISA 2015, Murillo y otros (2017) señalan que la segregación escolar de los estudiantes nacidos en el extranjero en el país es 0,45 (según el índice de Gorard). En la Cataluña, Valiente (2008) estima el valor de la segregación en el año 2001 en 0,55 (índice de Disimilitud) y 0,09 (índice de Aislamiento). Por su parte, el estudio del Síndic (2008) realizado con datos PISA 2006 estima su valor en 0,36 (valores medidos a través del índice de Disimilitud).

Como hemos reflejado en la revisión de la literatura, la investigación realizada hasta el momento tan sólo ofrece algunas evidencias de la segregación escolar en España en momentos puntuales. Por ello, la presente investigación busca conocer la evolución de la segregación escolar de carácter socioeconómico y por origen nacional en España y sus Comunidades Autónomas, de tal forma que se tengan evidencias del impacto de la crisis económica. Para ello, se hace una explotación especial de todos los datos disponibles del estudio PISA desde el año 2000 hasta 2015 (OCDE, 2016).

## Método

Para alcanzar este objetivo se analizaron los datos de PISA en sus ediciones 2000, 2003, 2006, 2009, 2012 y 2015, todas las disponibles

hasta el momento. El programa PISA estudia de forma trienal, desde el año 2000, los resultados de los sistemas educativos de diversos países en tres áreas fundamentales: Matemáticas, Lengua y Ciencias; y obtiene información de factores asociados al aprendizaje que permiten contextualizar y hasta cierto punto explicar los resultados obtenidos. Para la recogida de datos, PISA utiliza a parte de las pruebas de rendimiento, cuestionarios de contexto dirigidos a los directores de los centros, las familias y los estudiantes de Educación Secundaria.

La variable utilizada para medir la segregación escolar fue el índice de nivel socioeconómico y cultural NSEC (*Index of Economic, Social and Cultural Status*) definido por PISA. Este índice se compone de índices parciales relacionados con el contexto socioeconómico y cultural del estudiante (índice del mayor nivel educativo de los padres, índice del mayor nivel laboral de los padres, e índice del patrimonio cultural familiar). El índice NSEC para el total de los países que participan en PISA tiene un valor medio de cero y una desviación típica de uno. De manera que una puntuación negativa indica que las familias de los alumnos de un país tienen un nivel socioeconómico y cultural inferior a la media de los países participantes en PISA (OCDE, 2016).

## Muestra

Para la selección de la muestra, PISA utiliza un muestreo estratificado por conglomerados en dos etapas, con muestras en la primera etapa proporcionales al tamaño. Las unidades primarias de muestreo son las escuelas y las secundarias los alumnos de cada escuela, que se seleccionaron 35 de forma aleatoria a partir de listas de sujetos elegibles en cada escuela. Las variables de estratificación explícitas utilizadas para el caso español son la titularidad del centro, comunidad autónoma, tamaño de la escuela y modalidad de enseñanza en el País Vasco; las implícitas, los códigos postales y en Cataluña el tamaño de la localidad. La selección efectiva de la muestra se lleva a cabo mediante un procedimiento sistemático con un intervalo de muestreo. La distribución de estudiantes y escuelas de la muestra se encuentra en la Tabla II.

TABLA II. Muestra del estudio. Número de estudiantes, de centros y % de inmigrantes

	2000			2003			2006		
	Estud.	Centros	% Inmig	Estud.	Centros	% Inmig	Estud.	Centros	% Inmig
Andalucía							1463	51	2,10
Aragón							1526	51	4,95
Ppdo. Asturias							1.579	51	2,09
Illes Balears									
Canarias									
Cantabria							1.496	53	3,56
Castilla y León				1.490	51	1,46	1.512	52	2,74
Cast.-La Mancha									
Cataluña				1.515	50	3,47	1.527	51	7,91
Extremadura									
Galicia							1.573	53	1,54
La Rioja							1.333	45	5,81
C. de Madrid									
Región de Murcia									
C. F. de Navarra							1.590	52	6,93
País Vasco				3.885	141	1,48	3.929	151	3,49
C. Valenciana									
<b>Total España</b>	<b>6.135</b>	<b>185</b>	<b>2,54</b>	<b>10.790</b>	<b>383</b>	<b>2,83</b>	<b>19.604</b>	<b>684</b>	<b>6,10</b>

	2009			2012			2015		
	Estud.	Centros	% Inmig	Estud.	Centros	% Inmig	Estud.	Centros	% Inmig
Andalucía	1.416	51	5,20	1.434	52	3,06	1813	54	3,23
Aragón	1.514	52	11,52	1.393	49	12,36	1798	53	12,45
Ppdo. Asturias	1.536	54	4,59	1.611	55	5,86	1790	54	5,79
Illes Balears	1.463	52	13,00	1.435	54	16,60	1797	54	14,24
Canarias	1.448	50	9,62				1842	54	9,60
Cantabria	1.516	51	6,57	1.523	54	8,99	1924	56	9,00

<b>Castilla y León</b>	1.515	51	4,86	1.592	55	6,10	1858	57	6,60
<b>Cast.-La Mancha</b>							1889	55	7,63
<b>Cataluña</b>	1.381	50	9,49	1.435	51	11,12	1769	52	14,81
<b>Extremadura</b>				1.536	53	2,86	1809	53	1,74
<b>Galicia</b>	1.585	54	3,30	1.542	56	4,50	1865	59	5,07
<b>La Rioja</b>	1.288	46	12,05	1.532	49	16,76	1461	47	15,70
<b>C. de Madrid</b>	1.453	51	14,55	1.542	51	13,31	1808	51	14,84
<b>Reg. de Murcia</b>	1.321	51	11,49	1.374	51	14,85	1796	53	12,77
<b>C. F. de Navarra</b>	1.504	49	12,04	1.530	51	14,56	1874	52	13,10
<b>País Vasco</b>	4.768	177	4,35	4.738	174	8,17	3612	119	7,75
<b>C. Valenciana</b>	1.370	21					1625	53	12,28
<b>Total España</b>	25.887	889	8,42	25.312	893	7,85	6736	201	9,07

Nota: Las cifras sobre el porcentaje de inmigrantes están ponderadas.

Fuente: Elaboración propia a partir de PISA 2000, 2003, 2006, 2009, 2012 y 2015.

## Procedimiento

Para estimar la magnitud de la segregación escolar del conjunto de España, tanto para la segregación socioeconómica como por origen nacional, se utilizan cuatro índices de segregación diferentes: el índice de Gorard, el de Disimilitud, el de Hutchens y el de Aislamiento. Los tres primeros miran la dimensión de uniformidad de la segregación escolar, y el último la dimensión de exposición (Murillo, 2016). El sentido de utilizar tantas estrategias es aportar informaciones que sean comparables con estudios análogos, así como mostrar la estabilidad entre los diferentes resultados verificando con ello su fiabilidad. Las limitaciones de espacio de este artículo hacen que sea imposible mostrar los resultados de los cuatro índices en el análisis de la evolución por Comunidades; por ello se ha optado por utilizar el índice de Gorard dado que es el único que sus resultados se mantienen invariantes ante el tamaño del grupo minoritario, lo que le hace ser especialmente atractivo en el estudio de la evolución de la segregación (Gorard, 2009, Gorard y Taylor 2002). Un análisis de los índices puede ser encontrado en Murillo (2016).

El procedimiento seguido consiste en estimar la magnitud de la segregación socioeconómica, para cada año con datos disponibles, para el 25% de estudiantes con familias de menor NSEC (Q1) y para el 25% de estudiantes con familias de mayor NSEC (Q4). Esa doble mirada posibilita tener una panorámica más global que si solo se utiliza el Q1 como es habitual en este tipo de estudios. Para estimar la segregación escolar por origen nacional se considera como grupo minoritario a los estudiantes inmigrantes de primera generación, es decir, aquellos estudiantes no nacidos en España.

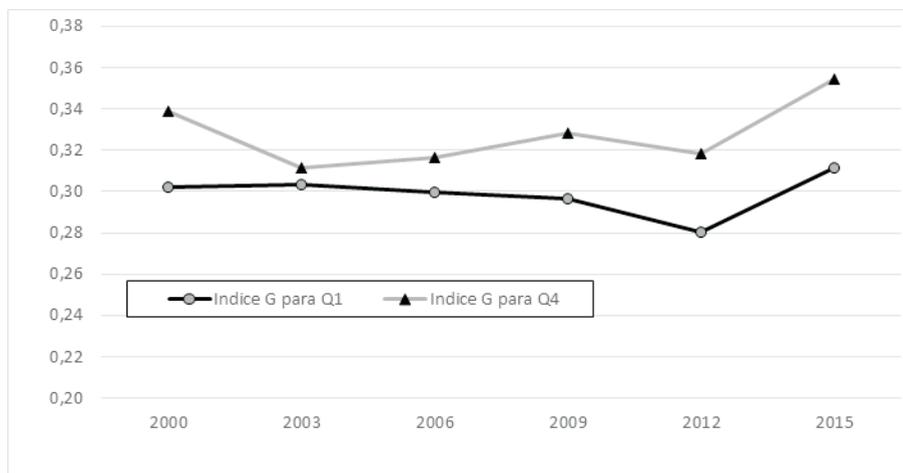
## Resultados

Acorde con los objetivos planteados, mostramos en primer lugar, los resultados del análisis de la evolución de la segregación por nivel socioeconómico, para centrarnos, posteriormente, en la segregación por origen nacional. En ambos casos, aportamos una visión global del conjunto de España y a continuación, una mirada a cada una de las 17 Comunidades Autónomas de forma independiente.

### Evolución de la segregación escolar por nivel socioeconómico

La segregación escolar por nivel socioeconómico en Educación Secundaria en España ha disminuido ligeramente desde 2000 hasta 2012, a partir de entonces ha sufrido un fuerte incremento (Gráfico I). Este hecho se observa independientemente del índice, y tanto si se considera el 25% de los estudiantes con familias de menor NSEC (Q1), como si se analiza la segregación con el 25% de los estudiantes de familias con mayor NSEC (Q4) (Tabla III).

**GRÁFICO I.** Evolución de la segregación escolar por nivel socioeconómico en Educación Secundaria Obligatoria en España. Índice de Gorard para el 25% de los estudiantes con familias de menor NSEC (Q1) y el 25% de mayor (Q4)



Fuente: Elaboración propia a partir de PISA 2000, 2003, 2006, 2009, 2012 y 2015.

Efectivamente, si nos centramos en la magnitud de la segregación según el índice de Gorard para el 25% de los estudiantes con familias de menor NSEC para el conjunto del Estado español disminuye de 0,30 en 2000 a 0,28 en 2012, y luego aumenta hasta los 0,31.

Estimar la segregación con cualquiera de los otros índices ofrece un perfil análogo. Así, con el índice de Disimilitud ha disminuido en 0,03 puntos desde el año 2000 y hasta el 2012 (Tabla III). En 2015, sin embargo, esa cifra sube de nuevo hasta 0,42 superando la cota alcanzada en 2000. El índice de Hutchens pasa de 0,16 en 2000 hasta 0,14 en 2012 y de ahí a 0,17; y, por último, el de Aislamiento de 0,39 baja a 0,37 en 2012, y a 0,39 en 2015. Teniendo en cuenta que los tres primeros índices estiman la segregación en su dimensión de uniformidad (aunque con postulados y resultados diferentes) y el de Aislamiento mide la dimensión de exposición, un comportamiento tan parecido da muestra de la fiabilidad de los resultados.

Estos datos podrían dar una respuesta tentativa y provisional a la pregunta de investigación que orienta este trabajo. Una plausible razón

del significativo incremento de la segregación desde 2012 a 2015 podría ser la crisis económica que vive España desde 2008. En las conclusiones retomaremos esta idea.

**TABLA III.** Evolución de la segregación escolar por nivel socioeconómico en España para el 25% de los estudiantes con familias de menor NSEC (Q1)

	2000	2003	2006	2009	2012	2015
<b>Índice de Gorard</b>	0,3023	0,3036	0,2997	0,2963	0,2806	0,3113
<b>Índice de Disimilitud</b>	0,4030	0,4045	0,3991	0,3945	0,3732	0,4152
<b>Índice de Raíz Cuadrada</b>	0,1617	0,1612	0,1563	0,1582	0,1434	0,1734
<b>Índice de Aislamiento</b>	0,3854	0,3898	0,3823	0,3788	0,3699	0,3892

Fuente: Elaboración propia a partir de PISA 2000, 2003, 2006, 2009, 2012 y 2015.

Una imagen complementaria nos la aporta el análisis de la magnitud de la segregación para el 25% de los estudiantes con familias de mayor NSEC (Q4). De entrada, la segregación escolar en España se ha mantenido significativamente más alta para los estudiantes con familias del Q4 que las de Q1.

En este caso, la disminución de la segregación entre 2000 y 2012 no es lineal, sino que, del 2000 a 2003 baja, posteriormente sube hasta el 2009, cuando se produjo una ligera bajada. El aumento entre el 2012 y 2015, en este caso, ha sido mucho más marcado para el Q4, que para el Q1. Efectivamente, para los cuatro índices utilizados la disminución de 2000 a 2012 es de 4,8% para el índice de Hutchens, del 4% para el de Aislamiento, de 2,6% para el de Disimilitud, y de 2,1% para el de Gorard. En el año 2015, se recupera lo perdido y en tan sólo tres años la segregación sube entre un 3,6% y un 4,69% dependiendo del índice utilizado (Tabla IV).

Así, es posible anticipar una segunda idea conclusiva de la posible incidencia de la crisis en la segregación escolar. Los datos indican que la crisis posiblemente está afectando a una mayor concentración de los estudiantes en función del NSEC de sus familias, pero esta concentración es mayor entre los estudiantes con familias de mayor NSEC que con los de menor.

**TABLA IV.** Evolución de la segregación escolar por nivel socioeconómico en España para el 25% de los estudiantes con familias de mayor NSEC (Q4)

	<b>2000</b>	<b>2003</b>	<b>2006</b>	<b>2009</b>	<b>2012</b>	<b>2015</b>
<b>Índice de Gorard</b>	0,3387	0,3116	0,3164	0,3281	0,3182	0,3542
<b>Índice de Disimilitud</b>	0,4514	0,4156	0,4227	0,4377	0,4253	0,4722
<b>Índice de Raíz Cuadrada</b>	0,1970	0,1528	0,1675	0,1707	0,1552	0,1954
<b>Índice de Aislamiento</b>	0,4493	0,4031	0,4186	0,4322	0,4097	0,4459

Fuente: Elaboración propia a partir de PISA 2000, 2003, 2006, 2009, 2012 y 2015.

El análisis detallado de la evolución de la magnitud de la segregación escolar por NSEC en cada Comunidad nos muestra una imagen llena de matices propios de sus características y de las diferentes medidas tomadas en estos 15 años analizados.

Comenzando el análisis por Comunidades para el 25% de los estudiantes con familias de menor NSEC (Tabla V), se observa que no en todas las comunidades ocurre el fenómeno descrito a nivel estatal:

- En Andalucía, Comunidad de Madrid y Cataluña se observa un incremento de la segregación de prácticamente ininterrumpida desde la fecha en la que se tienen los primeros datos (2003, 2006 o 2009) hasta 2015.
- En Castilla y León y Galicia, por el contrario, se percibe una bajada en la segregación en los años analizados, aunque en ambas con algún altibajo.
- El comportamiento de la segregación en La Rioja es igual al del conjunto de Estado: baja hasta el 2012, momento en el que sufre una fuerte subida.
- En Aragón y en las Illes Balears se produce un fenómeno inverso al del conjunto de España: sube hasta 2012 y desde ese momento se observa una marcada disminución.

En el resto, o no se observa una tendencia clara (Asturias, Cantabria, Murcia y País Vasco), o no hay datos suficientes para sacar conclusiones (Canarias, Castilla-La Mancha, Extremadura y Comunitat Valenciana).

**TABLA V.** Evolución de la segregación escolar de carácter socioeconómico en Educación Secundaria Obligatoria en España. Índice de Gorard para el 25% de los estudiantes con familias con menor NSEC (Q1)

	2000	2003	2006	2009	2012	2015
<b>Andalucía</b>			0,2578	0,2599	0,2599	0,2606
<b>Aragón</b>			0,2452	0,2608	0,2991	0,2221
<b>Ppdo. de Asturias</b>			0,2420	0,2775	0,2534	0,2636
<b>Illes Balears</b>				0,2489	0,2564	0,2019
<b>Canarias</b>				0,2681		0,2569
<b>Cantabria</b>			0,2334	0,2528	0,1998	0,2230
<b>Castilla y León</b>		0,2633	0,2366	0,2799	0,2493	0,2274
<b>Castilla-La Mancha</b>						0,2623
<b>Cataluña</b>		0,2824	0,2778	0,2979	0,2946	0,3047
<b>Extremadura</b>					0,2609	0,2254
<b>Galicia</b>			0,2875	0,2742	0,2900	0,2294
<b>La Rioja</b>			0,2552	0,2290	0,2189	0,2712
<b>Com. de Madrid</b>				0,2780	0,3326	0,3632
<b>Región de Murcia</b>				0,2745	0,2219	0,2438
<b>C. Foral de Navarra</b>			0,2701	0,2535	0,2467	0,2525
<b>País Vasco</b>		0,2752	0,2834	0,2732	0,2770	0,2678
<b>Com. Valenciana</b>						0,2234
<b>España</b>	0,3023	0,3036	0,2997	0,2963	0,2806	0,3113

Fuente. Elaboración propia a partir de PISA 2000, 2003, 2006, 2009, 2012 y 2015.

El estudio pormenorizado de la segregación escolar en la diferentes Comunidades para el 25% de estudiantes con familias de mayor NSEC ofrece una gran cantidad de matices que nos dan cuenta de las diferentes realidades existentes en cada Región producto, entre otras cosas, de las distintas políticas educativas. A modo de cata, analizamos cuatro regiones representativas, comenzando por la Comunidad de Madrid, que ostenta el dudoso honor de ser la más segregadora de España. La misma, tras bajar del 2009 al 2012, repunta su índice de nuevo en 2015, producto seguramente de la crisis, pero también de las políticas llevadas a cabo en estos últimos años. Cataluña, tras mantenerse el índice más o menos estable de 2003 a 2012, se incrementa a partir de entonces de forma llamativa. El País Vasco, sin embargo, desde 2003 a 2005 ha logrado

disminuir su segregación de una forma importante año tras año. Y, por último, en Andalucía no se observa una tendencia clara en ningún sentido en los 9 años de los que hay datos (Tabla VI).

**TABLA VI.** Evolución de la segregación escolar de carácter socioeconómico en Educación Secundaria Obligatoria en España. Índice de Gorard para el 25% de los estudiantes con familias con mayor NSEC (Q4)

	2000	2003	2006	2009	2012	2015
<b>Andalucía</b>			0,2970	0,2893	0,3040	0,2885
<b>Aragón</b>			0,2933	0,3113	0,3122	0,2569
<b>Ppdo. de Asturias</b>			0,3320	0,3097	0,3215	0,3392
<b>Illes Balears</b>				0,3191	0,2375	0,2396
<b>Canarias</b>				0,2747		0,3447
<b>Cantabria</b>			0,2915	0,2951	0,2600	0,2775
<b>Castilla y León</b>		0,3023	0,2412	0,2886	0,2756	0,2831
<b>Castilla-La Mancha</b>						0,3443
<b>Cataluña</b>		0,3008	0,3362	0,2898	0,3088	0,3466
<b>Extremadura</b>					0,2829	0,2689
<b>Galicia</b>			0,3325	0,2986	0,2975	0,2647
<b>La Rioja</b>			0,2672	0,2778	0,2777	0,2505
<b>Com. de Madrid</b>				0,4027	0,3751	0,3925
<b>Región de Murcia</b>				0,3091	0,2727	0,2843
<b>C. Foral de Navarra</b>			0,3044	0,2916	0,3152	0,2931
<b>País Vasco</b>		0,3131	0,3129	0,2887	0,2825	0,2794
<b>Com. Valenciana</b>						0,2793
<b>España</b>	0,3387	0,3116	0,3164	0,3282	0,3182	0,3542

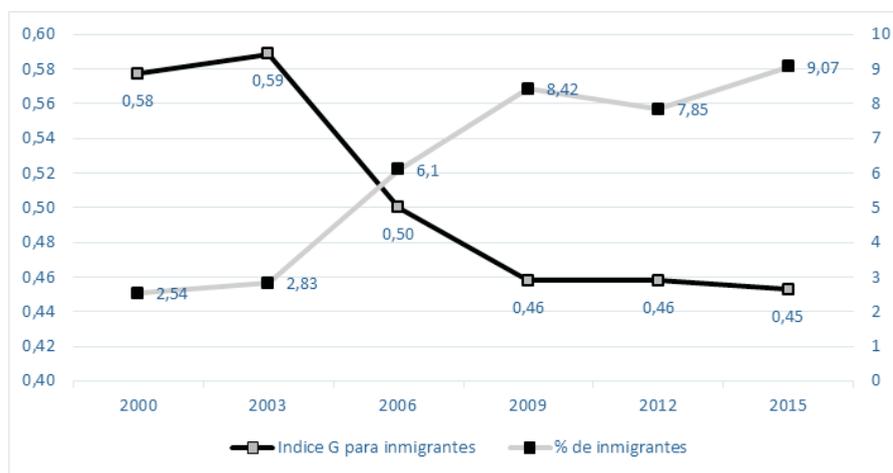
Fuente. Elaboración propia a partir de PISA 2000, 2003, 2006, 2009, 2012 y 2015.

## Evolución de la segregación escolar por origen nacional

Una imagen complementaria nos la aporta el estudio de la evolución de la distribución de los estudiantes en los centros educativos en función de si han nacido en España o en el extranjero (inmigrantes de primera generación).

Dada la estrecha relación entre la proporción de estudiantes del grupo minoritario y la segregación (Murillo, 2016), el Gráfico II refleja la evolución de la magnitud de la segregación y del porcentaje de extranjeros matriculados. Los resultados indican que el porcentaje de estudiantes inmigrantes ha experimentado un crecimiento prácticamente ininterrumpido desde el año 2000 (2,54%) a 2015 (9,07%) y, a la par, que la segregación ha disminuido, de 0,58 en 2000 a 0,45 en 2015. También es destacable que esa disminución de la segregación se paró en 2009 y se mantuvo prácticamente estable hasta 2015, a pesar del aumento del porcentaje de inmigrantes esos años.

**GRÁFICO II.** Evolución del porcentaje de estudiantes extranjeros escolarizados en Educación Secundaria Obligatoria en España, y de la segregación escolar por origen nacional. Índice de Gorard para inmigrantes de primera generación



Fuente. Elaboración propia a partir de PISA 2000, 2003, 2006, 2009, 2012 y 2015.

Aunque el índice de Gorard es el más adecuado para los propósitos evolutivos dado que es el único que no se ve afectado por el tamaño del grupo minoritario (Gorard, 2009; Murillo, 2016) es posible observar que este comportamiento se valida con matices con el cálculo del resto de índices. El comportamiento de los tres índices que miden la

dimensión de uniformidad es muy similar (con mínimas diferencias); mientras que los resultados del índice de Aislamiento, que mide la dimensión de exposición, son radicalmente distintos. Este hecho se explica comprendiendo su significado: el índice de Aislamiento estima la probabilidad de que un estudiante extranjero se encuentre en su centro con otro de su misma condición. Así, es un índice muy influido por el tamaño relativo del grupo minoritario... cuantos más extranjeros hay mayor será la probabilidad de encontrarse ante ellos (Tabla VII).

La idea conclusiva que podemos extraer del análisis es que la posible incidencia de la crisis ha sido menor para la segregación escolar por origen. Ciertamente es que se ha parado la tendencia a la baja desde 2000, pero sus efectos no han sido tan dramáticos como para la segregación socioeconómica.

**TABLA VII.** Evolución de la segregación escolar por origen nacional en España para inmigrantes de primera generación

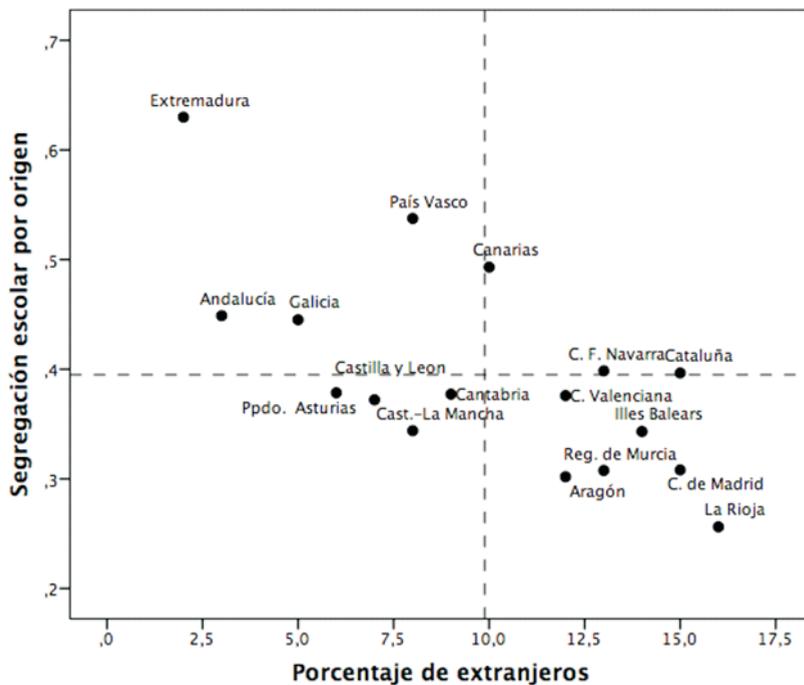
	2000	2003	2006	2009	2012	2015
<b>Índice de Gorard</b>	0,5775	0,5888	0,5002	0,4579	0,4580	0,4529
<b>Índice de Disimilitud</b>	0,5926	0,6060	0,5328	0,5000	0,5000	0,4980
<b>Índice de Raíz Cuadrada</b>	0,4127	0,4152	0,3313	0,2697	0,2771	0,2535
<b>Índice de Aislamiento</b>	0,1723	0,1169	0,2241	0,2102	0,2271	0,2296

Fuente. Elaboración propia a partir de PISA 2000, 2003, 2006, 2009, 2012 y 2015.

El análisis de la evolución de la segregación escolar por origen en cada Comunidad de nuevo ofrece un panorama lleno de matices tan rico que resulta difícil de describir en un espacio tan limitado como es este formato (Tabla VIII). En todo caso, dos ideas llaman la atención.

En primer lugar, es de destacar, de nuevo, la gran variabilidad entre la segregación de las diferentes Comunidades, que no es del todo explicada por la diferente proporción de alumnado extranjero. Así, aunque la correlación entre segregación y porcentaje de inmigrante es alta ( $\rho = -0,629$ , sig. = 0,007), se observa que algunas Comunidades tienen una segregación mucho mayor de lo esperado (Gráfico III).

GRAFICO III. Relación entre el porcentaje de extranjeros y el índice de segregación escolar por origen nacional de Gorard



Fuente. Elaboración propia a partir de PISA 2015.

La tendencia evolutiva de la segregación escolar por origen en cada Comunidad es, de nuevo, muy diferente. Así, aunque en todas ellas (con la cautela que genera el no contar la serie completa en todos los casos) la tendencia es la disminución de la segregación, en cinco Comunidades se observa un incremento entre el año 2012 y 2015, en nueve se observa una disminución, y no hay datos para sacar conclusiones en las tres restantes. Las cinco Comunidades en las que aumenta la segregación en estos últimos años, en orden de magnitud del incremento, son Extremadura, Navarra, Cantabria, Murcia e Illes Balears. Al contrario, las Comunidades en donde más baja la segregación escolar por origen del 2012 al 2015 son, y en este orden: Castilla y León, Madrid, Andalucía, Aragón y País Vasco. Una de las explicaciones a estas variaciones se encuentra en los

cambios en el porcentaje de estudiantes extranjeros, de tal forma que mayor un incremento en el porcentaje de estudiantes inmigrantes está asociado a una bajada en el índice de segregación de Gorard ( $\rho = 0,613$ , sig. = 0,020).

**TABLA VIII.** Evolución de la Segregación escolar por origen en España. Índice de Gorard para inmigrantes de primera generación por CCAA

	2000	2003	2006	2009	2012	2015
<b>Andalucía</b>			0,6791	0,5333	0,5224	0,4488
<b>Aragón</b>			0,4466	0,3070	0,3452	0,3019
<b>Ppdo. de Asturias</b>			0,5907	0,4585	0,4079	0,3785
<b>Illes Balears</b>				0,3901	0,3332	0,3432
<b>Canarias</b>				0,5045		0,4931
<b>Cantabria</b>			0,5118	0,3923	0,3660	0,3772
<b>Castilla y León</b>		0,7049	0,6364	0,4365	0,4914	0,3721
<b>Castilla-La Mancha</b>						0,3439
<b>Cataluña</b>		0,5312	0,3721	0,3981	0,3975	0,3966
<b>Extremadura</b>					0,4987	0,6299
<b>Galicia</b>			0,7716	0,4005	0,4604	0,4451
<b>La Rioja</b>			0,3818	0,3465	0,2612	0,2562
<b>Com. de Madrid</b>				0,3329	0,4004	0,3081
<b>Región de Murcia</b>				0,3623	0,2974	0,3076
<b>C. Foral de Navarra</b>			0,4415	0,4215	0,3243	0,3985
<b>País Vasco</b>		0,7892	0,7486	0,6075	0,5757	0,5375
<b>Com. Valenciana</b>						0,3759
<b>España</b>	0,5775	0,5888	0,5002	0,4579	0,4580	0,4529

Fuente: Elaboración propia a partir de PISA 2000, 2003, 2006, 2009, 2012 y 2015.

## Discusión y Conclusiones

Esta investigación ha mostrado con claridad que el esfuerzo realizado desde hace 15 años por reducir la segregación escolar en España se ha visto bruscamente interrumpido por las circunstancias sociales y económicas que han caracterizado el panorama nacional de crisis en

los últimos años. Tanto si consideramos la segregación socioeconómica, como por origen nacional, los datos son persistentes en mostrar un claro cambio de tendencia los últimos años, al menos para el total del Estado español, de una progresiva disminución de la segregación escolar desde 2000 hasta 2012 a un fuerte incremento en estos tres últimos años; la explicación más plausible es la repercusión tanto en la sociedad como en el sistema educativo de una fuerte recesión económica comenzada en el año 2007 y cuya remisión está aún lejos de percibirse. El hecho de que PISA analice a estudiantes de 15 años –es decir, que accedieron al centro educativo en el que están matriculados hace al menos tres o cuatro años–, puede explicar que haya un efecto retardado entre la incidencia de la crisis en las familias y la segregación escolar.

Esta incidencia parece ser mayor en la segregación de los estudiantes con familias de mayor NSEC, algo menor entre los estudiantes de menor NSEC y una ruptura de la tendencia a la baja en la segregación por origen. Es decir, los datos parecen indicar que los estudiantes con familias de menores recursos han abandonado las escuelas en las que predominan los estudiantes de familias con mayores ingresos y viceversa, lo que fomenta la existencia de escuelas para ricos y escuelas para pobres, y ha llevado a que España sea el sexto país con mayor segregación de la Unión Europea (Murillo y Martínez-Garrido, 2018).

Resulta muy arriesgado establecer relaciones de causa-efecto sin una mínima reflexión al respecto. Así, la existencia de tendencias diversas en las diferentes Comunidades nos lleva a pensar en la posibilidad de la influencia de otros factores. En la Comunidad de Madrid, por ejemplo, la región con mayor segregación escolar socioeconómica, políticas como el establecimiento del distrito único, la generalización de centros bilingües, el crecimiento de la red de centros privados o la publicación, más o menos soterrada, de rankings escolares en una clara tendencia hacia un sistema de cuasi-mercado escolar, podrían explicar el incremento de la segregación. Sin embargo, aunque no contamos aun con estudios sobre la incidencia de crisis en la segregación escolar, trabajos análogos sobre la segregación residencial de carácter socioeconómica han encontrado que la misma se ha incrementado de forma importante en los años de la crisis, tanto en diversos países (p. ej., Andersson y Hedman, 2016; Zwiers, Bolt, Van Ham y Van Kempen, 2016), como en España (Bayona-i-Carrasco, Gil-Alonso, Rubiales-Pérez y Pujadas-Rúbies, 2017; Leal y Sorando, 2016). Dada la estrecha relación entre ambos fenómenos (Lareau y Goyette,

2014), todo hace pensar que la crisis juega un papel fundamental en ese claro crecimiento. En todo caso, son necesarias más investigaciones para profundizar en esta hipótesis.

La segregación escolar atenta directamente contra la igualdad de oportunidades y, con ello, genera una sociedad menos cohesionada y desigual. El carácter multicausal de la segregación (Gorard y Fitz, 2000) hace que la lucha contra ella sea especialmente compleja. Así, aunque hay factores de difícil modificación a corto plazo –como la segregación residencial antes comentada–, existen políticas educativas que pueden incrementar o reducir la segregación. Los recortes a la educación pública en forma de disminución en el número de docentes, especialistas y de sus condiciones laborales o de calidad de las instalaciones, el incremento de estudiantes por grupo o su repercusión en el deterioro, deliberado o no, de la imagen social de los centros de titularidad pública en detrimento de los privados, por ejemplo, generan un desplazamiento de los que pueden elegir hacia algunos centros que se convierte en segregación. Si, de verdad, la crisis económica ha acabado, es preciso que se tomen decisiones políticas de compensación que hagan que se recupere la tendencia a la baja en la segregación.

No podemos obviar en esta investigación las limitaciones que para el estudio de la segregación escolar tiene el uso de las bases de datos de PISA. Pero antes de su listado, es preciso señalar que PISA es la única fuente de información de la que se dispone que permite no solo hacer un estudio evolutivo (al realizarse cada tres años con la misma metodología) como este, sino tener una estimación de la segregación escolar mínimamente fiable. No olvidemos que se analizan datos de 94.464 estudiantes escolarizados en 3.235 centros a lo largo de los 15 años del estudio PISA. Una limitación, sin duda es que contar con pocos estudiantes por escuela podría ocultar segregación entre aulas, así como otro tipo de segregaciones intra-escolares.

Esta investigación es apenas un primer diagnóstico de una preocupante realidad muy poco abordada en España, pero solo eso. De ahí que sea necesario el desarrollo de futuros estudios que ayuden a identificar qué políticas educativas favorecen la desagregación escolar, que profundicen en el papel de los centros privados en la segregación, que exploren la incidencia de la segregación en el rendimiento académico de los y las estudiantes, así como en el desarrollo de valores de convivencia o que contribuyan a una mejor comprensión de los criterios de selección de escuela por parte de las familias y la incidencia de la crisis, entre otros.

Un sistema educativo segregado lleva irremediablemente a una sociedad desigual e injusta. La crisis económica ha logrado que España se convierta en uno de los países más desiguales de Europa. Si de verdad la crisis ha acabado, es preciso tomar medidas urgentes que reviertan la situación de una segregación escolar creciente. En caso contrario, el sueño hacia una sociedad más justa e inclusiva se irá inexorablemente alejando.

### Referencias bibliográficas

- Alegre, M. A., Benito, R. y González-Baetbó, I. (2008). Procesos de segregación y polarización escolar: La incidencia de las políticas de zonificación escolar. *Profesorado. Revista de Currículum y Formación del Profesorado*, 12(2), 1-26.
- Andersson, R. y Hedman, L. (2016). Economic decline and residential segregation: A Swedish study with focus on Malmö. *Urban Geography*, 37(5), 748-768. <https://doi.org/10.1080/02723638.2015.1133993>
- Ayllón, S. (2015). *Infancia, pobreza y crisis económica*. Barcelona: Obra Social “La Caixa”
- Bayona-i-Carrasco, J., Gil-Alonso, F., Rubiales-Pérez, M. y Pujadas-Rúbies, I. (2017). New spatial mobility patterns in large Spanish cities: From the economic boom to the great recession. *Applied Spatial Analysis and Policy*, 10, 1-26. <https://doi.org/10.1007/s12061-017-9222-x>
- Benito, R. y González-Baetbó, I. (2007). *Processos de segregació escolar a Catalunya*. Barcelona: Mediterrània.
- Bonal, X. (2012a). *Municipis contra la segregació escolar. Sis experiències de política educativa local*. Barcelona: Fundació Jaume Bofil.
- Bonal, X. (2012b). Education policy and school segregation of migrant students in Catalonia: The politics of non-decision-making. *Journal of Education Policy*, 27(3), 401-421. <https://doi.org/10.1080/02680939.2011.645168>
- Cebolla-Boado, H. y Garrido Medina, L. (2011). The impact of immigrant concentration in Spanish schools: School, class, and composition effects. *European Sociological Review*, 27(5), 606-623. <https://doi.org/10.1093/esr/jcq024>

- Coleman, J., Hoffer, T. y Kilgore, S. (1982). *High school achievement: Public, catholic and private schools compared*. Nueva York, NY: Basic Books.
- Duncan, O. y Duncan, B. (1955). A methodological analysis of segregation indexes. *American Sociological Review*, 20, 210-217. <https://doi.org/10.2307/2088328>
- Ferrer, F. (Dir.). (2008). *Les desigualtats educatives a Catalunya*. Barcelona: Fundació Jaume Bofill.
- Ferrer, F., Castejón, A., Castel, J. L. y Zancajo, A. (2011). *PISA 2009: Avaluació de les desigualtats educatives a Catalunya*. Barcelona: Fundació Jaume Bofill.
- García-Gómez, S. y Cabanillas, M. C. (2017). Incidencia de la crisis económica en el alumnado de educación primaria. La perspectiva del profesorado. *Revista Complutense de Educación*, 28(1), 185-201. [https://doi.org/10.5209/rev\\_RCED.2017.v28.n1.49075](https://doi.org/10.5209/rev_RCED.2017.v28.n1.49075)
- Gorard, S. (2009). Does the index of segregation matter? The composition of secondary schools in England since 1996. *British Educational Research Journal*, 35(4), 639-652. <https://doi.org/10.1080/01411920802642389>
- Gorard, S. y Fitz, J. (2000). Investigating the determinants of segregation between schools. *Research Papers in Education*, 15(2), 115-132. <https://doi.org/10.1080/026715200402452>
- Gorard, S. y Taylor, C. (2002). What is segregation? A comparison of measures in terms of 'strong' and 'weak' compositional invariance. *Sociology*, 36(4), 875-895. <https://doi.org/10.1177/003803850203600405>
- Gorard, S., Hordosy, R. y See, B. H. (2013). Narrowing down the determinants of between-school segregation: An analysis of the intake to all schools in England, 1989-2011. *Journal of School Choice*, 7(2), 182-195. <https://doi.org/10.1080/15582159.2013.791182>
- Hutchens, R. (2004). One measure of segregation. *International Economic Review*, 45(2), 555-578. <https://doi.org/10.1111/j.1468-2354.2004.00136.x>
- Lareau, A. y Goyette, K. (2014). *Choosing homes, choosing schools*. Nueva York, NY: Russell Sage Foundation.
- Leal, J. y Sorando, D. (2016). Economic crisis, social change and segregation processes in Madrid. En T. Tammaru, S. Marciniak, M. Van Ham y S. Musterd (Eds.), *Socio-economic segregation in European capital cities: East meets west*. Londres: Routledge.

- Lieberson, S. (1981). An asymmetrical approach to segregation. En C. Peach (Ed.), *Ethnic segregation in cities* (pp. 61-83). Londres: Croom-Helm.
- Mancebón-Torrubia, M. J. y Pérez-Ximénez, D. (2009). Segregación escolar en el sistema educativo español. Un análisis a partir de PISA 2006. *Investigaciones de Economía de la Educación*, 4, 63-77.
- Mancebón-Torrubia, M. J. y Pérez-Ximénez, D. (2010). Una valoración del grado de segregación socioeconómica existente en el sistema educativo español. Un análisis por comunidades autónomas a partir de PISA 2006. *Regional and Sectoral Economic Studies*, 10(3), 129-148.
- Mas Ivars, M. Pérez García, F. y Uriel, E. (2015). *Capital público en España: Evolución y distribución territorial (1900-2012)*. Bilbao: Fundación BBVA.
- Murillo, F. J. (2016). Midiendo la segregación escolar en América Latina. Un análisis metodológico utilizando el TERCE. *REICE. Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, 14(4), 33-60. <https://doi.org/10.15366/reice2016.14.4.002>
- Murillo, F. J. y Martínez-Garrido, C. (2018). Magnitud de la Segregación escolar por nivel socioeconómico en España y sus Comunidades Autónomas y comparación con los países de la Unión Europea. *RASE: Revista de la Asociación de Sociología de la Educación*, 11(1), 37-58. <https://doi.org/10.7203/RASE.11.1.10129>
- Murillo, F. J., Martínez-Garrido, C. y Belavi, G. (2017). Segregación escolar por origen en España. *OBETS*, 12(2), 395-423. <https://doi.org/10.14198/OBETS2017.12.2.04>
- OCDE. (2016). *Programme of international student assessment (PISA 2015)*. París: OCDE.
- Orfield, G., Ee, J., Frankenberg, E. y Siegel-Hawley, G. (2016). *Brown at 62: School segregation by race, poverty and state*. Los Angeles, CA: Civil Rights Project, UCLA
- Owens, A., Reardon, S. F. y Jencks, C. (2016). Income segregation between schools and school districts. *American Educational Research Journal*, 53(4), 1159-1197. <https://doi.org/10.3102/0002831216652722>
- Reardon, S. F. y Owens, A. (2014). 60 Years after Brown: Trends and consequences of school segregation. *Annual Review of Sociology*, 40, 199-218. <https://doi.org/10.1146/annurev-soc-071913-043152>

- Sánchez Hugalde, A. (2007). La segregació escolar dels immigrants a Catalunya. *Quaderns D'avaluació*, 13, 55-76.
- Síndic. (2008). *La segregación escolar en Cataluña*. Barcelona: Síndic.
- Stephan, W. (Ed.). (2013). *School desegregation: Past, present, and future*. Nueva York, NY: Springer.
- Stroub, K. J. y Richards, M. P. (2013). From resegregation to reintegration: Trends in the racial/ethnic segregation of metropolitan public schools, 1993-2009. *American Educational Research Journal*, 50(3), 497-531. <https://doi.org/10.3102/0002831213478462>
- Valiente, O. (2008). ¿A qué juega la concertada? La segregación escolar del alumnado inmigrante en Cataluña (2001-06). *Profesorado. Revista de Currículum y Formación del Profesorado*, 12(2), 1-23.
- Van-Zanten, A. (2007). School differentiation and segregation in the Parisian periphery: An analysis of urban schools' logics of action and their effects. En W. Pink y G. Noblit (Eds.), *International handbook of urban education* (pp. 431-445). Dordrecht: Springer.
- Zwiers, M., Bolt, G., Van Ham, M. y Van Kempen, R. (2016). The global financial crisis and neighborhood decline. *Urban Geography*, 37(5), 664-684. <https://doi.org/10.1080/02723638.2015.1101251>

**Información de contacto:** F. Javier Murillo. Universidad Autónoma de Madrid, Facultad de Formación de Profesorado y Educación. Departamento de Didáctica y Teoría de la Educación. Universidad Autónoma de Madrid. C/ Tomas y Valente 3. Madrid E-28049, España. E-mail: javier.murillo@uam.es



# Análisis de las herramientas de medición de los Estilos de Aprendizaje<sup>1</sup>

## Analysis of the Learning Styles measurement tools

DOI: 10.4438/1988-592X-RE-2017-381-382

María Luz Diago Egaña  
María José Cuetos Revuelta  
Patricia González González

*Universidad Internacional de la Rioja*

### Resumen

La primera vez que aparece el constructo “Estilo de Aprendizaje” (Gibson, 1969), lo hace con un significado que denota el modo en que una persona prefiere que le presenten la información para llevar a cabo el aprendizaje. Esta definición ha ido evolucionando a lo largo de los años, al mismo tiempo que se desarrollaban numerosas herramientas para detectar cada una de sus dimensiones.

En este estudio se pretende elaborar un listado actualizado de herramientas disponibles para la detección de los distintos estilos de aprendizaje; analizar sus características métricas, la relación entre ellos y ponderar los cuestionarios en función de su utilización.

Para ello se realiza una búsqueda bibliográfica exhaustiva en distintos buscadores como Google académico, Web of Science (WOS), Dialnet Education Resources Information Center (ERIC), ScienceDirect y un análisis bibliométrico en el periodo comprendido entre el 21 de mayo de 2015 al 30 de abril de 2017.

Mostramos un listado actualizado de las herramientas disponibles para detectar los estilos de aprendizaje, basándonos en los cuatro criterios mínimos postulados por Coffield, Moseley, Hall y Ecclestone (2004). Se analizan las

---

<sup>(1)</sup> Este trabajo ha sido financiado por la Universidad Internacional de la Rioja (UNIR). Se enmarca dentro del proyecto de investigación titulado “Estilos de Aprendizaje aplicados a las Ciencias Experimentales” concedido por UNIR tras la convocatoria interna de proyectos de investigación de julio de 2016.

características métricas de los cuestionarios recogidos por ellos juntamente con otros encontrados en nuestra investigación, considerando también si dichos cuestionarios están o no validados.

Mientras que, el cuestionario CSI (Allison y Hayes, 1996) se presenta como el más robusto, si nos ceñimos a los cuestionarios que encajarían con la semántica original del término “Estilos de Aprendizaje” (Gibson, 1969) los que se revelan con más presencia en diferentes motores de búsqueda serían en VAK/VARK (Google Scholar en cualquier idioma), CHAEA (Google Scholar en español) e ILS (Felder y Soloman, 1997) (Web of Science y ScienceDirect). Se concluye que la ponderación porcentual de los cuestionarios basándonos en su aparición según diferentes buscadores, varía en función de la lengua de búsqueda y el buscador utilizado.

*Palabras clave:* Estilos de aprendizaje, canales de percepción, estrategias de aprendizaje, preferencias de aprendizaje, estilos cognitivos, neurociencia, cuestionario.

### **Abstract**

The first time the “Learning Style” construct (Gibson, 1969) appears, it does so with a meaning that denotes how a person prefers to be presented with the information to carry out the learning. This definition has evolved over the years, while numerous tools were developed to detect each of its dimensions. This study aims to develop an updated list of tools available for the detection of different learning styles; analyzing its metric characteristics and to weigh the questionnaires according to their use, as well as to analyze the relationship between the measurement tools of the learning styles.

A comprehensive bibliographic search is carried out by using different internet searching engines such as Google Scholar, Web of Science (WOS), Dialnet, Education Resources Information Center (ERIC), ScienceDirect and a bibliometric analysis performed in the period from May 21, 2015 to April 30, 2017.

We present an updated list of the tools available to detect learning styles, based on the four minimum criteria postulated by Coffield, Moseley, Hall y Ecclestone (2004). We analyze the metric characteristics of the questionnaires collected by them, together with others found in our research, also considering whether or not these questionnaires are validated.

While the CSI questionnaire (Allison and Hayes, 1996) is presented as the most robust, if we stick to questionnaires that would fit with the original semantics of the term “Learning Styles” (Gibson, 1969), inventories VAK/VARK (Google Scholar in any language), CHAEA (Google Scholar in Spanish) and ILS (Felder y Soloman, 1997) (Web of Science and ScienceDirect) would be the most widely used. It is concluded that the percentage weight of the questionnaires, based on their presence in literature, varies according to the search language and the search engine used.

*Keywords:* Learning styles, perception channels, learning strategies, learning preferences, learning styles, cognitive styles, neuroscience, questionnaire.

## Introducción

Los Estilos de Aprendizaje (EA) tienen una gran influencia dentro del campo de la educación y se han estudiado durante años, desde niveles que van desde el preescolar a la universidad. Uno de los principales objetivos del estudio y determinación de los EA ha sido, durante años, mejorar en general los resultados -inmediatos y a largo plazo- del proceso de enseñanza-aprendizaje.

De hecho, durante los últimos 40-50 años se ha creado una floreciente industria en torno al concepto de EA, dedicada a publicar tanto test como guías para maestros, en la cual muchas organizaciones ofrecen talleres de desarrollo profesional para profesores y educadores en torno a este concepto (Pashler, McDaniel, Rohrer y Bjork, 2008). En todas estas intervenciones la idea fundamental que subyace detrás de este constructo es que cada uno de nosotros tiene un estilo específico de aprendizaje (a veces llamada “preferencia”), y aprendemos mejor cuando la información se nos presenta en este estilo.

La literatura referente a los EA es amplia y controvertida, empezando por la propia definición de este constructo. En general, los defensores de la teoría de los estilos de aprendizaje sostienen que la instrucción óptima para los alumnos requiere el diagnóstico del estilo de aprendizaje de los individuos y, en consecuencia, la adaptación de la instrucción a los mismos. Y conocer los estilos de aprendizaje que utilizan los alumnos permite que el docente sepa qué metodología utilizar y cómo llevarla a cabo de una forma más efectiva. Hay autores que determinan que está suficientemente probado que los estudiantes aprenden con más efectividad cuando se les enseña con sus EA predominantes (Alonso, Gallego y Honey, 1999). Sin embargo, según Alonso et ál. (1999) no sólo hay que tener en cuenta el EA de los alumnos sino también el Estilo de Enseñar de los profesores. En este sentido existen estudios que señalan la necesidad de analizar los EA de los alumnos y relacionarlos con los estilos de enseñanza aplicados por los profesores (Coloma, Manrique, Revilla y Tafur, 2008; Pupo, 2012; González-Peiteado, 2013; Aiello, García, y Jaramillo, 2015), ya que esto puede facilitar la posibilidad de éxito académico de los alumnos (Saarikoski, Salojärvi, Del Corso y Ovcin, 2001). Sin embargo, Pashler et ál. (2008) encontraron que, a pesar de que la literatura sobre los estilos de aprendizaje es enorme, prácticamente no hay ninguna evidencia que apoye la idea de que si la instrucción se

proporciona en un formato que coincide con la preferencia del alumno, se mejora el rendimiento académico del mismo. Estos autores apuntan a que el diseño de la mayor parte de las investigaciones es débil, y que incluso investigaciones con un diseño experimental eficaz encontraron resultados que contradicen categóricamente las populares suposiciones acerca de los estilos de aprendizaje. De hecho, algunas de las debilidades se correlacionan con el diseño de las propias herramientas para medir estilos de aprendizaje.

A lo largo de los años se han realizado numerosos intentos de determinar los EA de los estudiantes a través del diseño de cuestionarios. A medida que ha ido evolucionado la definición del constructo, lo han hecho también numerosas herramientas de medición para detectar cada una de las dimensiones de los EA. Así encontramos cuestionarios que miden diversos parámetros como la personalidad (MSP de Apter, Mallows y Williams, 1998; MBTI (Myers y McCaulley, 1985); la interacción dentro del grupo (SLSQ de Grasha, 1996); el nivel cognitivo (LSI de Kolb, 1984); el nivel emocional (CHAEA de Alonso, Gallego y Honey, 1994); el modo de procesar la información (hemisferio izquierdo o derecho, el HBDI de Herrmann, 1982); el ambiente (Dunn, Dunn y Price, 1989); el canal de percepción de la información (VARK de Fleming y Mills 1992), el tiempo de respuesta (MFFT de Kagan, 1966); el contexto de aprendizaje (cuestionario Grasha, 1996); las preferencias básicas en relación con las metas, las actitudes y los sentimientos (cuestionario LIFO de Atkins, citado en García, Santizo y Alonso, 2009), entre otros.

A la vista de lo anteriormente expuesto queda patente el elevado número de herramientas disponibles para determinar distintos estilos de aprendizaje, basados en múltiples aspectos y examinados en distintas etapas educativas, predominando los estudios realizados en la Educación Superior frente a la Educación Primaria y Secundaria. Por otra parte, estos modelos no han estado exentos de críticas, sobre todo debido a la escasa fiabilidad y validez que según Curry (1990) y, Coffield et ál. (2004), poseen las mediciones realizadas.

## Objetivos de estudio

En este trabajo se plantea primeramente elaborar un listado de herramientas disponibles actualmente para la detección de los distintos

Estilos de Aprendizaje y analizar sus características métricas, y en segundo lugar analizar cuáles son los cuestionarios más utilizados por los distintos autores para medir EA y la relación existente entre las herramientas de medición.

## **Método**

La metodología empleada en este trabajo se basa en una exhaustiva revisión bibliográfica en las siguientes bases de datos: Scholar Google, Web of Science (WOS), Dialnet, ERIC y ScienceDirect.

## **Procedimiento**

Este trabajo parte de la selección de los 13 cuestionarios que Coffield et ál. establecen como principales en 2004. Se realizó una exhaustiva revisión bibliográfica desde el 21 de mayo de 2015 al 30 de abril de 2017 empleando como criterio de búsqueda cada una de las palabras clave que figuran en la Tabla I (Anexo), tomadas individualmente o combinaciones de estas. Tras la lectura del gran volumen de artículos encontrados se filtraron aquellos que nos remitían a instrumentos de medición. Los criterios de selección de dichos instrumentos fueron incluir aquellos que medían estilos de aprendizaje y que, o bien no fueron considerados originalmente por Coffield et ál (2004), o si fueron considerados no figuran como principales, o bien son posteriores a dicho trabajo. Asimismo se tuvo en cuenta que el constructo de estos instrumentos fuera acorde a la definición de EA de Gibson (1969). Una vez seleccionados de esta manera los instrumentos se utilizaron sus nombres y acrónimos (Tabla I), para realizar el análisis de cada instrumento. Los criterios de análisis fueron determinar el constructo, las dimensiones, comprobar si el instrumento había sido validado o no, quién lo ha validado y las características psicométricas.

El análisis bibliométrico acerca de cuáles son los cuestionarios más utilizados en los últimos años se realizó a través de los buscadores de GoogleScholar, Web of Science y ScienceDirect, tanto en castellano como en inglés, teniendo en cuenta los siguientes criterios:

- Nombre completo del instrumento en cuestión
- Siglas del instrumento+apellido/s de autor/es

El promedio de estas búsquedas es el dato cuantitativo resultante que se presenta en los diferentes gráficos del apartado de resultados.

## Resultados y Discusión

### Herramientas de detección de EA

Desde que en la década de 1950 se empezó a hablar de EA son muchas las herramientas que se han diseñado con el fin de detectarlos. En 2004, Coffield et ál. realizan el análisis de 71 cuestionarios para medir EA que existían hasta la fecha, categorizando 13 como principales y concluyendo que, de estos 13, únicamente el Cognitive Style Index (CSI) (Allinson y Hayes, 1996) reúne cuatro criterios mínimos esenciales para que el cuestionario pueda ser utilizado en investigaciones sucesivas (consistencia interna, fiabilidad test-retest, validez de constructo y validez predictiva). Por el contrario los cuestionarios de Riding (1991) (CSA: análisis de estilos cognitivos de Riding), de Sternberg (1999) (TSI, cuestionario de estilos de Pensamiento de Sternberg) no reúnen ninguno de los criterios mínimos esenciales y por tanto no se van a considerar en este trabajo.

La Tabla II (Anexo) refleja información sobre el CSI (Coffield et ál. 2004).

En el trabajo de Coffield et ál. (2004) se señala que además dos cuestionarios, el MSP (Motivational Style Profile) de Apter, Mallows y Willians (1998) y el ILS (Inventory of Learning Styles) de Vermunt (1996) reúnen 3 de los 4 criterios mínimos necesarios en un cuestionario. El cuestionario de Apter et ál. (1998) reúne los criterios de consistencia interna, fiabilidad test-retest, y validez predictiva, mientras que el cuestionario de Vermunt (1996) reúne los criterios de consistencia interna, fiabilidad test-retest, y validez del constructo (Tabla III, anexo).

Asimismo en el trabajo de Coffield et ál. (2004) se señala que 3 cuestionarios más reúnen 2 de los 4 criterios mínimos necesarios, los cuestionarios ASSIST (Approaches and Study Skills Inventory for Students) de Entwistle (1988), HDBI (Herrmann's Brain Dominance Instrument) de Herrmann (1989) y el MBTI (Myers-Briggs Type Indicator) (Myer y McCaulley, 1985, 1998).

Finalmente Coffield et ál. (2004) señala que 4 cuestionarios más reúnen 1 de los 4 criterios mínimos necesarios, tal y como se muestra en la Tabla IV (Anexo). A mayores, el cuestionario LSP (Jackson's Learning Styles Profiler) de Jackson (2002) (Tabla IV) que inicialmente no fue validado por Coffield et ál. (2004) por falta de datos, ha mostrado a posteriori validez en 1 de los cuatro criterios (Tabla IV).

Por tanto, y como se puede comprobar en la Tabla III, de las herramientas validadas por Coffield et ál. (2004) el cuestionario CSI mide **preferencias procesar la información**, los cuestionarios HDBI, LSQ y LSI miden **preferencias de aprendizaje**, mientras que los cuestionarios VAK/VARK y GSD miden **modalidades en la percepción** de la información y **canales específicos para procesar** la información respectivamente. Estas herramientas serían las relacionadas con nuestro constructo.

A parte de estos cuestionarios analizados en Coffield et ál. (2004), se han buscado otros cuestionarios utilizados para medir EA. La información obtenida al respecto se puede consultar en la Tabla V (Anexo), dónde se muestran las características métricas de los mismos.

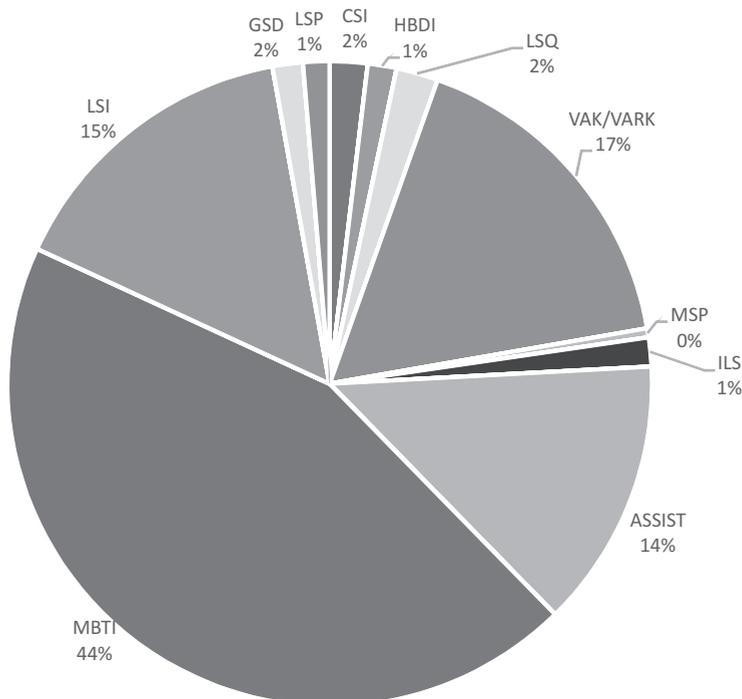
Como puede observarse en dicha Tabla V, los cuestionarios que sirven para medir preferencias de aprendizaje sólo son el Test de Barsch, ILS (Felder-Soloman), LSS y CESEA.

A continuación, analizamos cuáles de estos cuestionarios que miden preferencias de aprendizaje, canales de percepción o estilos cognitivos han seguido utilizándose en los últimos años. La Tabla VI (Anexo) muestra los trabajos más recientes que han empleado, en la metodología del trabajo de campo, estos instrumentos.

Como complemento a la Tabla VI y con el objetivo de estudiar la repercusión que han tenido la utilización de los 11 cuestionarios que según Coffield et ál. (2004) presentan de 1-4 criterios mínimos para ser validados, se ha efectuado una búsqueda a través de Google Académico utilizando para ello los nombres de los distintos cuestionarios y acotando la búsqueda desde que se publicó su trabajo, por lo tanto al periodo 2004-2016, y en toda la web. Los resultados de dicha búsqueda se reflejan en el gráfico I.

**GRÁFICO I.** Resultados de búsqueda de los diferentes cuestionarios en cualquier idioma (n=10167 resultados obtenidos).

**PROMEDIO RESULTADOS BÚSQUEDAS 11 CUESTIONARIOS  
COFFIELD EN CUALQUIER IDIOMA (Google Scholar)**



Fuente: elaboración propia

Se puede observar en este Gráfico I que el CSI, sobre el que Coffield et ál. (2004) establecía que se reunían los cuatro criterios mínimos esenciales para que el cuestionario pudiera ser utilizado en investigaciones posteriores, sólo ha sido encontrado en el 2% de los trabajos localizados. Por el contrario, y con mucha diferencia, es el cuestionario MBTI (que mide características de la personalidad y estructura cognitiva) al que hacen referencia un 44% de los trabajos encontrados en Google Académico para el periodo en cuestión.

Si nos fijamos únicamente en los cuestionarios relacionados con nuestro constructo, este gráfico nos indica que son los cuestionarios LSI y VAK/VARK los que reflejan un mayor número de trabajos localizados, en concreto un 15% y 17% respectivamente. Señalar que aunque el cuestionario LSI mida preferencias de EA, reúne solo uno de los 4 criterios marcados por Coffield et ál. (2004) para ser validado, y no fue validado por autores como Reynolds (2003) y Shum (2003). Asimismo el cuestionario VAK/VARK mide modalidades en la percepción de la información, y al igual que el anterior sólo reúne un criterio para ser validado (Coffield et ál. 2004), pero aunque a posteriori ha sido validado por Leite, Svinicki y Shi (2010), estos aconsejan tener precaución en su utilización en investigación en lo referente a la redacción de los ítems y la puntuación de los algoritmos de la escala. Además, y según Fleming (2012), técnicamente VAK/VARK no se refiere a EA, pues este cuestionario sólo proporciona información sobre las modalidades preferidas de comunicación.

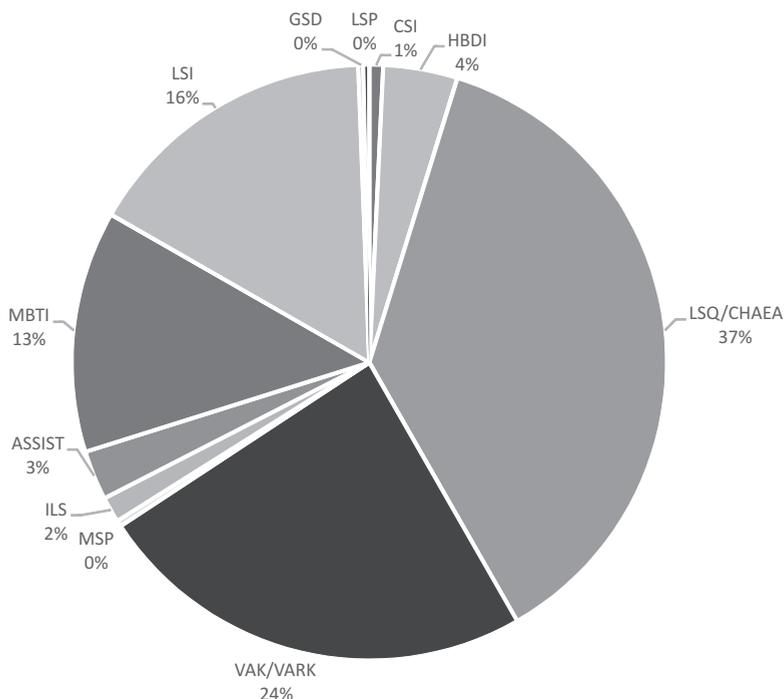
Al proceder de la misma manera, pero únicamente realizando la búsqueda en idioma español, el resultado obtenido se puede analizar en el gráfico II.

Se aprecia en el gráfico II que en español, el CSI incluso aparece en un menor porcentaje que en cualquier idioma, mientras que cuestionario CHAEA (adaptación al castellano del LSQ de Honey y Mumford (1982) aumenta considerablemente su presencia, pasando de un 2% a un 37%, seguido del VAK/VARK con un 24%. Recordemos que para ambos cuestionarios, Coffield et ál. (2004), sólo encuentran un criterio mínimo esencial para ser utilizados.

Si a estos cuestionarios analizados por dichos autores añadimos los reflejados en la Tabla V, los resultados obtenidos se muestran a continuación. Así, el gráfico III muestra, para el mismo periodo 2004-2016, como queda la presencia de resultados en toda la web cuando consideramos todos los cuestionarios, tanto los analizados por Coffield et ál. (2004) como los estudiados en dicha tabla. En esta búsqueda se ha omitido el LSS porque este cuestionario fue diseñado en 2014 y hasta 2016 sólo ha sido utilizado por los autores que lo crearon.

**GRÁFICO II.** Resultados de búsqueda de los diferentes cuestionarios sólo en idioma español. (n=930 resultados obtenidos).

PROMEDIO RESULTADOS BÚSQUEDAS 11 CUESTIONARIOS  
COFFIELD EN ESPAÑOL (Google Scholar)



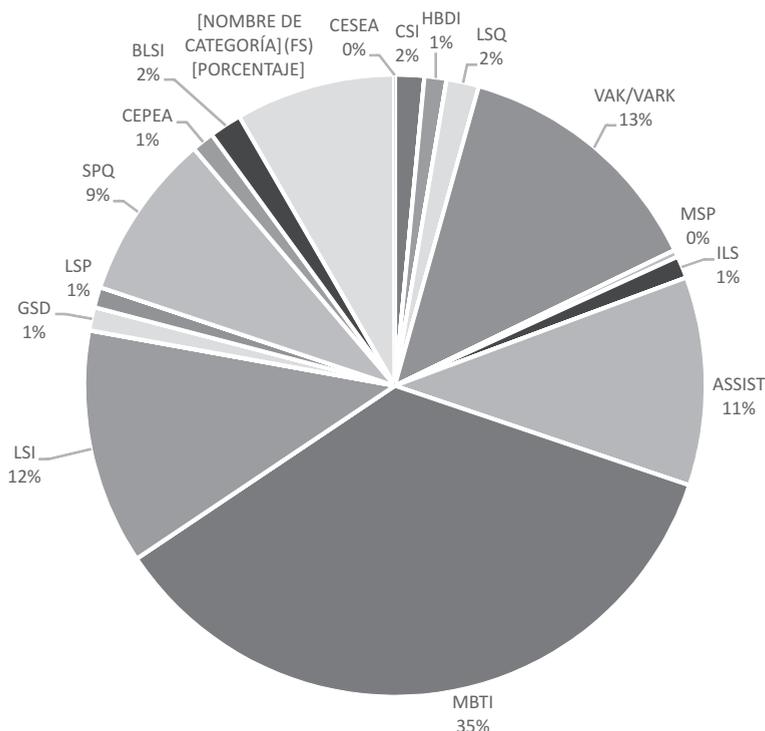
Fuente: elaboración propia

Según el gráfico III, los cuestionarios mayormente utilizados son MBTI, LSI, VAK/VARK y ASSIST, mientras que los cuestionarios SPQ e ILS (Felder-Soloman, se ha marcado como FS en el gráfico) aparecen en este análisis con una presencia ligeramente significativa, del 9% y 8% respectivamente.

De las herramientas relacionadas con nuestro constructo, el LSI no está validado (Coffield et ál. sólo encuentran 1 criterio de los 4 necesarios), el VAK/VARK, como hemos comentado anteriormente, aunque está validado en 2010 por Leite et ál., se aconseja utilizarlo con precaución.

**GRÁFICO III.** Resultados de búsqueda de todos los cuestionarios en toda la web. (n=12688 resultados obtenidos).

**PROMEDIO RESULTADOS BÚSQUEDAS TODOS LOS CUESTIONARIOS EN CUALQUIER IDIOMA (Google Scholar)**



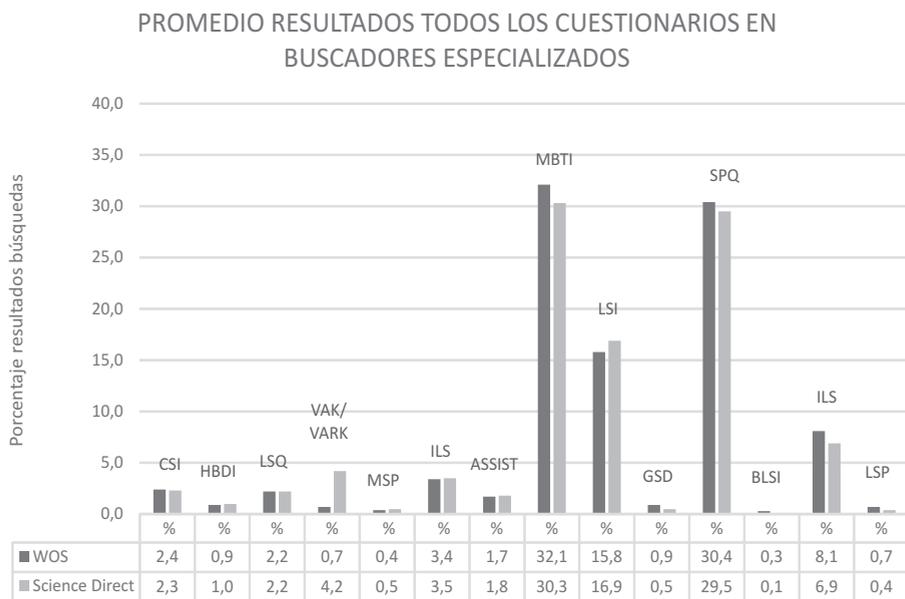
Fuente: elaboración propia

Es el cuestionario ILS (Felder-Soloman) el que mide preferencias o tendencias de aprendizaje y ha sido validado recientemente (Zywno, 2003; Felder y Spurlin, 2005; Hosford y Siders, 2010; Wang y Mendori, 2015; por lo que podría ser empleado como herramienta para medir preferencias de aprendizaje.

Con la idea de ser más rigurosos en este estudio y con el objetivo de ampliarlo, las búsquedas hechas en Google Académico se han llevado a

cabo también en buscadores especializados, en concreto se han realizado en Web of Science (WOS) y ScienceDirect. La única diferencia con el estudio anterior es que no se ha diferenciado por idioma. El resultado obtenido se puede ver en el gráfico IV.

**GRÁFICO IV.** Resultados de búsqueda de todos los cuestionarios en buscadores especializados. (WOS n=837 resultados obtenidos; ScienceDirect n=1046 resultados obtenidos).



Fuente: elaboración propia

En el gráfico IV no aparecen los datos el cuestionario CEPEA porque no se ha encontrado ningún registro en estos buscadores. Tampoco se muestran los datos sobre CESEA porque sólo se han encontrado 2 registros en WOS y ninguno en ScienceDirect. Si nos fijamos únicamente en los cuestionarios relacionados con nuestro constructo, los que aparecen con una mayor presencia y en orden son el LSI y el ILS (Felder-Soloman).

El LSI aunque mide preferencias de EA Experimental, no fue validado por Reynolds (2003) y Shum (2003) y si por Coffield et ál. (2004), quien

indica que solo reúne 1 de los 4 criterios. De hecho, se han seguido desarrollando diferentes versiones del cuestionario LSI a lo largo de los años siendo el más reciente el KLSI versión 4 (2013) pero sólo ha sido validado por los propios autores del cuestionario.

Por tanto, la búsqueda en otras bases de datos –WOS y ScienceDirect– nos conduce de nuevo al cuestionario ILS (Felder-Soloman) como candidato para ser utilizado en la medición de EA ya que mide preferencias o tendencias de aprendizaje, está entre los más utilizados y está validado externamente y en fechas recientes. Asimismo, aunque el cuestionario CSI ha resultado ser de los menos utilizados durante los últimos años, es el único que reúne los cuatro criterios marcados por Coffield et ál. (2004) y ha sido también validado recientemente por Armstrong y Qi (2016).

No obstante, para ambos cuestionarios también encontramos detractores. Así para el cuestionario ILS (Felder-Soloman), Al-Azawei, Parslow y Lundqvist (2015) consideran en general este cuestionario válido para diagnosticar EA pero señalan la moderada solidez de alguno de los criterios analizados (confiabilidad y validez de percepción) o baja consistencia interna de algunos de los constructos. En la misma línea Çardak y Selvi (2016) ponen de manifiesto la necesidad de realizar más estudios que confirmen la validez de este cuestionario ya que los resultados de los test de validación difieren según las culturas a las que pertenezcan los estudiantes, hecho que ya había sido previamente apuntado por Joy y Kolb (2009).

Por otro lado, para el cuestionario CSI, aunque las pruebas de confiabilidad interna y confiabilidad prueba-retest han mostrado fuertes resultados (Coffield et ál. 2004), otros investigadores, han criticado la estructura de factores subyacentes del cuestionario inicial. Mientras que para algunos su naturaleza es unifactorial (Allinson y Hayes, 1996; Armstrong y Qi, 2016), otros proponen una naturaleza claramente multifactorial, donde las dos dimensiones, intuición y análisis, han de tratarse de manera separada pero correlacionadas (Hodgkinson y Sadler-Smith, 2003; Backhaus y Liff, 2007).

Backhaus y Liff en su artículo del 2007, analizando la naturaleza de dicha estructura, aplica este cuestionario a una muestra de 222 estudiantes universitarios de negocios estadounidenses y encuentra diferencias en lo obtenido si lo compara con lo que observaron Allinson y Hayes en 1996 en su estudio británico. Una de las posibles razones a estas diferencias la

atribuyen a la redacción de los ítems del cuestionario, pues su lectura por nativos estadounidenses puede dar lugar a interpretaciones diferentes ya que están acostumbrados a una versión diferente del inglés. Por los resultados que obtienen, los autores concluyen que es posible que este instrumento no haya podido medir con precisión cómo procesa la información ésta muestra en concreto, estableciendo que creen necesario revisar la semántica del cuestionario para eliminar así cualquier disparidad en el uso del idioma. Esta conclusión también coincide con lo expuesto anteriormente por Hodgkinson y Sadler-Smith (2003), que establecieron que el proceso de puntuación y los problemas de redacción del cuestionario CSI dan lugar a resultados insatisfactorios y que necesitan ser revisados. Así mismo para que los resultados obtenidos al aplicar este cuestionario sean lo más representativos posible, la muestra a estudiar deber ser amplia y lo más heterogénea posible (Backhaus y Liff, 2007).

## Relación entre algunas de las herramientas utilizadas para medir EA

A continuación mostramos la relación entre algunas de las herramientas que, a lo largo de los años se han venido utilizando para la detección de EA:

**ILS** (Felder-Soloman) tiene una versión adaptada al castellano por Martínez-Fernández, García-Ravidá, González-Velázquez, Gutiérrez-Braojos, Poggioli, Ramírez-Otalvaro y Tellería (2009). Este cuestionario trata de identificar las estrategias, las motivaciones y las actitudes frente al estudio o las tareas de aprendizaje. El cuestionario se compone de dos partes, la primera dedicada a identificar dentro de una lista de 55 actividades aquellas que los estudiantes realizan en el contexto de sus estudios, y la segunda parte (24 ítems) dedicada a analizar los motivos, objetivos y actitudes que los estudiantes tienen con respecto a sus estudios.

**LSQ** se adaptó al español pasando a denominarse CHAEA (cuestionario Honey-Alonso de EA). Estos autores dividen los EA en cuatro categorías: activo, reflexivo, teórico y pragmático. Baus (2007) amplía y complementa aún más estas definiciones, estableciendo que, en el estilo activo, la persona improvisa, arriesga, descubre y es espontánea, por lo que es un alumno dinámico en clase. El estilo reflexivo se refiere a personas receptivas, analíticas y observadoras. El estilo pragmático se caracteriza por que los

alumnos son rápidos, decididos, planificadores, concretos, con objetivos definidos y seguros. Y, por último, las personas que muestran un estilo de teórico se caracterizan por disciplinadas, sistemáticas, ordenadas, sintéticas, razonadoras, pensadoras, perfeccionistas, y buscadoras de modelos teóricos.

Dunn, Dunn y Price (1975) en un inicio determinaron tres modalidades, que denominaron canales de percepción: visual, auditivo y kinestésico (**VAK**). Posteriormente, Fleming y Mills (1992) aumentaron las modalidades añadiendo la “lectura” (reading) (**VAR**K). No es fácil determinar al consultar la literatura al respecto, si a estos canales se les considera o no estilos de aprendizaje (Gamboa-Mora, Briceño-Martínez y Camacho-González, 2015). Según Fleming y Baume (2006), y sólo desde un punto de vista teórico, a estas modalidades sólo se las podría considerar como una parte para ser incluida dentro de los Estilos de aprendizaje. Por otra parte, Fleming (2012), opina que técnicamente VARK no es un Estilo de aprendizaje, sino que se refiere a cómo las personas aprenden, y se centra en las diferentes modalidades en las que las personas preferirían aprender. Y por lo tanto dicho cuestionario sólo aportaría información sobre modalidades preferidas de comunicación y sus resultados serían únicamente indicativos, no diagnósticos. Que una persona muestre una alta preferencia por uno de los canales no significa que el resto de modalidades no existan en esa misma persona.

La relación entre los enfoques de CHAEA y VAK, está determinada en función del contenido que se quiera enseñar en clase (Quiñones, 2004). Según este autor, las estrategias de enseñanza que se determinen deben tener en cuenta los diferentes canales de percepción: visual, auditivo y kinestésico, pues estos actuarían como precursores en el desarrollo de los diferentes estilos de aprendizaje (activo, reflexivo, teórico y pragmático).

La corriente de la Programación Neurolingüística (PNL) así como Dunn, Dunn y Price (1985), establece que las personas perciben el mundo desde los tres canales de percepción ya citados: visual, auditivo y kinestésico (VAK), pero añade que los individuos, en su gran mayoría, tienden a desarrollar alguno más que otro, llegando a presentar un canal perceptivo líder. Por lo que se puede concluir que la teoría en el VAK concuerda con los desarrollos de la Programación Neurolingüística (PNL), en cambio la del CHAEA no (Gamboa et ál. 2015).

Por otra parte Escanero-Marcén, Soria, Guerra-Sánchez y Silva (2016) concluyen que CHAEA y el cuestionario de Felder-Silvermann no son

efectivos del todo, por lo que elaboran un nuevo cuestionario, basado como los anteriores en el aprendizaje experiencial de Kolb y en los estilos cognitivos de Allison y Hayes (1996). El nuevo cuestionario se llama: Escanero-Soria de estilos de aprendizaje (CESEA). A la hora de preparar y seleccionar los ítems que integran el CESEA los autores tuvieron en cuenta el modelo cognitivo del Cognitive Style Index (CSI). Según Escanero-Marcén et ál. (2016) la equivalencia entre los estilos de CHAEA y los polos del CESEA sería la siguiente: Los cuatro estilos que explora el CHAEA, y que se representan sobre los cuatro ejes del diagrama, son activo, reflexivo, teórico y pragmático, se corresponden con los polos dinámico, reflexivo, teórico y operativo, respectivamente, del CESEA. En el CESEA, en cambio, los estilos son los cuadrantes y tienen las propiedades de los dos polos que los delimitan.

## Conclusiones

Las conclusiones de este trabajo son las siguientes:

De los 15 cuestionarios analizados en este artículo únicamente miden preferencias de aprendizaje, canales de percepción o estilos cognitivos los cuestionarios CSI, HDBI, LSQ, LSI, VAK/VARK, GSD, Test de Barsch, ILS (Felder-Soloman), LSS y CESEA, por lo que son estos los que consideramos relacionados con nuestro constructo. De ellos sólo el CSI cumple los cuatro criterios (consistencia interna, fiabilidad test-retest, validez de constructo y validez predictiva) según Coffield et ál. (2004).

Al analizar la presencia/relevancia en la red de estos cuestionarios, lo que encontramos es que, por una parte, existen algunas diferencias en cuanto a la relevancia de cada artículo si buscamos trabajos en según qué idioma (o en toda la web o sólo en castellano) o si buscamos en buscadores más o menos especializados (Google Scholar o WOS y ScienceDirect).

Independientemente del idioma o del buscador utilizado, la presencia/relevancia en la red del cuestionario CSI únicamente está entre el 1-2.5%, pese a estar validado por diferentes autores. Aún así otros critican su estructura, el proceso de puntuación y señalan problemas de redacción de los ítems.

En el resto de cuestionarios analizados por Coffield et ál. (2004) relacionados directamente con los EA, encontramos que en toda la web

los que aparecen con una mayor presencia son el VAK/VARK y LSI. Mientras que si atendemos sólo a trabajos en castellano, el cuestionario más relevante con diferencia es el CHAEA (adaptación al castellano del LSQ de Honey y Mumford, 1982), seguido del VAK/VARK.

Si nos fijamos en todos los cuestionarios analizados en este artículo relacionados directamente con los EA pero teniendo en cuenta el tipo de buscador, encontramos que en ambos tipos de buscadores aparecen como relevantes el LSI y el ILS (Felder-Soloman). Encontrando que la única diferencia entre ambos tipos reside en que, en el buscador genérico el VAR/VARK aparece como relevante, y esto no ocurre cuando buscamos en los buscadores específicos.

De estos cuatro cuestionarios más relevantes: LSI, LSQ, VAK/VARK e ILS (Felder-Soloman), sólo los dos últimos han sido validados por diferentes autores. El cuestionario ILS (Felder-Soloman) mide preferencia o tendencias de aprendizaje mientras que el cuestionario VAK/VARK mide modalidades en la percepción de la información. La Tabla VII (Anexo) muestra un resumen de las características de ambos cuestionarios. Cualquiera de las dos herramientas encajaría con la semántica original del término “Estilos de Aprendizaje” (Gibson, 1969).

La limitación encontrada en este trabajo es centrar la revisión bibliográfica exclusivamente en dos idiomas, castellano e inglés, aunque consideramos que engloban la mayor parte de la literatura sobre este tema.

La prospectiva inmediata del mismo sería la validación del cuestionario ILS en el idioma castellano con el objetivo de utilizarlo para determinar EA en alumnos de primaria y secundaria y poder seguir profundizando sobre las estrategias, técnicas y métodos de enseñanza que se adapten a dichos EA.

## Referencias bibliográficas

Abdollahimohammad, A. y Ja'afar, R. (2014). Learning style preferences of nursing students at two universities in Iran and Malaysia. *Journal of educational evaluation for health professions*, 11, 30, 1-5. doi: 10.3352/jeehp.2014.11.30

- Abdollahimohammad, A. y Ja'afar, R. (2015). Associations of learning style with cultural values and demographics in nursing students in Iran and Malaysia. *Journal of educational evaluation for health professions*, 12, 42, 1-6. doi: 10.3352/jeehp.2015.12.42
- Aiello, R. C., García, M. R. y Jaramillo, M. (2015). Determinación de los estilos de aprendizaje de estudiantes de 1er curso de Ingeniería Industrial y Electrónica de la Universidad Técnica del Norte. Ibarra. Ecuador: *Journal of Learning Styles*, 7, 14, 43-67.
- Al Hamdani, D. (2015). Exploring students' learning style at a Gulf University: a contributing factor to effective instruction. *Procedia-Social and Behavioral Sciences*, 176, 124-128. doi: 10.1016/j.sbspro.2015.01.452
- Alavi, S. S. y Makarem, J. (2016). Learning Style and Attitude toward Computer among Iranian Medical Students. *Journal of Medical Education*, 14, 1, 20-25.
- Al-Azawei, A., Parslow, P. y Lundqvist, K. (2015). A Psychometric Analysis of Reliability and Validity of the Index of Learning Styles (ILS). *International Journal of Psychological Studies*, 7, 3, 46-57.
- Aljaberi, N. M. (2015). University Students' Learning Styles and Their Ability to Solve Mathematical Problems. *International Journal of Business and Social Science*, 6, 4, 152-165.
- Allinson, C. W. y Hayes, J. (1996). The Cognitive Style Index: A Measure of Intuition-Analysis For Organizational Research. *Journal of Management Studies*, 33, 119-135. doi: 10.1111/j.1467-6486.1996.tb00801.x
- Alonso, C. M., Gallego, D. J. y Honey, P. (1994). *Cuestionario Alonso-Honey de Estilos de Aprendizaje (CHAEA)*. Recuperado de [https://www.researchgate.net/publication/265265933\\_CUESTIONARIO\\_HONEY-ALONSO\\_DE\\_ESTILOS\\_DE\\_APRENDIZAJE\\_CHAEA](https://www.researchgate.net/publication/265265933_CUESTIONARIO_HONEY-ALONSO_DE_ESTILOS_DE_APRENDIZAJE_CHAEA)
- Alonso, C. M., Gallego, D. J. y Honey, P. (1999). *Los estilos de aprendizaje. Procedimientos de diagnóstico y mejora*. Bilbao: Ediciones Mensajero.
- Apter, M. J., Mallows, R. y Williams, S. (1998). The development of the Motivational Style Profile. *Personality and Individual Differences*, 24, 1, 7-18.
- Armstrong, S. J. y Qi, M. (2016). A reassessment of the factor structure of the Allison-Hayes Cognitive Style Index. *Personality and Individual Differences*, 101, 240-242.

- Backhaus, K. y Liff, J. P. (2007). Cognitive Style Index: Further investigation of the factor structure with an American student sample. *Educational Psychology*, 27, 1, 21-31.
- Barbosa, S. D., Gerhardt, M. W. y Kickul, J. R. (2007). The role of cognitive style and risk preference on entrepreneurial self-efficacy and entrepreneurial intentions. *Journal of Leadership and Organizational Studies*, 13, 4, 86-104.
- Barca, A. (1999). *Manual del Cuestionario de Procesos de Estudio y Aprendizaje (CEPEA)*. A Coruña: Publicaciones de la Revista Galego-Portuguesa de Psicoloxía e Educación.
- Barch, J. (1996). *Learning styles inventory*. Leveroni Court. Novato, USA: Academic Therapy Publications.
- Barros, D. M. V. (2010). Estilos de uso do espaço virtual: Novas perspectivas para os ambientes de aprendizagem online. *Journal of Learning Styles*, 3, 6, 103-127.
- Baus, T. (2007). *Los estilos de aprendizaje*. Recuperado de <http://www.monografias.com/trabajos12/loestils/loestils.shtml>.
- Bhagat, A., Vyas, R. y Singh, T. (2015). Students awareness of learning styles and their perceptions to a mixed method approach for learning. *International Journal of Applied and Basic Medical Research*, 5, 1, S58-S65. doi: 10.4103/2229-516X.162281
- Biggs, J. (1985). Test review of Barsch Learning Style Inventory. En: J. V. Mitchell (Ed.), *The ninth mental measurements yearbook* (3-4). "Recuperado de <http://marketplace.unl.edu/buros/>"
- Biggs, J. (1987). *Reflective thinking and school learning: An introduction to the theory and practice of metacognition* (Item IO, Set 2. 14). Wellington, NZ: NZCER.
- Bothma, T. J. y de Boer, A. L. (2015). *Reflecting on Diverse Teaching Methodologies for an Information Literacy Programme for Large Groups*. The Third European Conference on Information Literacy (ECIL). Tallinn, Estonia, 19-22, octubre.
- Bunderson, C. V. (1994). The validity of the Herrmann Brain dominance instrument. En N. Herrmann (Ed.), *The creative brain* (364). Lake Lure, NC: Brain Books.
- Burnett, P. C. y Dart, B. C. (2000). The study process questionnaire: a construct validation study. *Assessment and Evaluation in Higher Education*, 25, 1, 93-99. doi: 10.1080/713611415

- Çardak, Ç. S. y Selvi, K. (2016). The Construct Validity of Felder-Soloman Index of Learning Styles (ILS) for the Prospective Teachers. *Gaziantep University Journal of Social Sciences*, 15, 2, 675-693.
- Cheng, G. (2014). Exploring students' learning styles in relation to their acceptance and attitudes towards using Second Life in education: A case study in Hong Kong. *Computers and Education*, 70, 105-115.
- Coffield, F., Moseley, D., Hall, E. y Ecclestone, K. (2004). *Learning styles and pedagogy in post-16 learning. A systematic and critical review*. Londres: Learning & Skills Research Centre. "Recuperado de <http://www.leerbeleving.nl/wp-content/uploads/2011/09/learning-styles.pdf>"
- Coloma, C. R., Manrique, L., Revilla, D. y Tafur, R. (2008). Estudio descriptivo de los estilos de aprendizaje de docentes universitarios. *Revista de Estilos de Aprendizaje*, 1, 1, 124-142.
- Curry, L. (1990). One critique of the research on learning styles. *Educational Leadership*, 48, 50-56.
- De Boer, A. L., Bothma, T. y du Toit, P. (2011). Enhancing information literacy through the application of whole brain strategies. *Libri*, 61, 1, 67-75. doi: 10.1515/libr.2011.006
- Díaz, M. A. C. y Aly, J. L. M. (2015). Estilos de aprendizaje en estudiantes universitarios: recursos informáticos como estrategia para su evaluación. *UNAH INNOV@*, 4, 33-39.
- Donche, V., Coertjens, L. y Van Petegem, P. (2010). Learning pattern development throughout higher education: A longitudinal study. *Learning and Individual differences*, 20, 3, 256-259.
- Dunn R., Dunn K. y Price G. (1975). *Learning Style Inventory (LSI)*. Lawrence, Kansas: Price System.
- Dunn R., Dunn, K. y Price G. (1979). *Learning Style Inventory (LSI) for Students in grades 3-12*. Lawrence, Kansas: Price Systems
- Dunn, R. y Dunn, K. (1978). *Teaching students through their individual learning styles: A practical approach*. New Jersey: Prentice Hall.
- Dunn, R., Dunn, K. y Price, G. (1985). *Manual: Learning Style Inventory*. Lawrence, Kansas: Price Systems.
- Dunn, R., Dunn, K. y Price, G. (1989). *Learning Style Inventory (LSI) for Students in grades 3-12*. Lawrence, Kansas: Price Systems.
- Entwistle, N. (1988). *Styles of learning and teaching*. London: David Fulton.

- Escanero-Marcén, J. F., Soria, M. S., Guerra-Sánchez, M., y Silva, J. (2016). Comparación de los estilos de aprendizaje de los alumnos de medicina obtenidos con un nuevo cuestionario con los proporcionados por el cuestionario Honey-Alonso (CHAEA). *FEM: Revista de la Fundación Educación Médica*, 19, 1, 19-26.
- Felder, R. M. y Spurlin, J. (2005). Applications, reliability and validity of the index of learning styles. *International Journal of Engineering Education*, 21, 1, 103-112.
- Felder, R. M. y Soloman, B. A. (1997). *Index of Learning Styles*. “Recuperado de <https://www.webtools.ncsu.edu/learningstyles/>”.
- Fleming, N. D. (2001). *Teaching and learning styles: VARK strategies*. Christchurch, New Zealand: N.D. Fleming.
- Fleming, N. D. y Mills, C. (1992). *Not another inventory, rather a catalyst for reflection*. “Recuperado de <http://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1245&context=podimproveacad>”
- Fleming, N. y Baume, D. (2006). Learning Styles Again: VARKing up the right tree! *Educational Developments*, SEDA Ltd, 7, 4, 4-7.
- Fleming, N. D. (2012). *Facts, Fallacies and Myths: VARK and Learning Preferences*. “Recuperado de <http://vark-learn.com/wp-content/uploads/2014/08/Some-Facts-About-VARK.pdf>”.
- Gamboa-Mora M. C., Briceño-Martínez J. J. y Camacho-González J. P. (2015). Caracterización de estilos de aprendizaje y canales de percepción de estudiantes universitarios, 31, 3, 509-527. “Recuperado de <http://www.redalyc.org/pdf/310/31045567026.pdf>”
- García, J. L, Santizo, J. A. y Alonso, C. M. (2009). Instrumentos de medición de estilos de aprendizaje. *Review of learning styles*, 4, 29, 3-21.
- Gargallo, B., Almerich, G., Suárez, J. M., García, E. y Garfella, P. R. (2012). Learning styles and approaches to learning in excellent and average first-year university students. *European journal of psychology of education*. 28, 4, 1361-1379. doi: 10.1007/s10212-012-0170-1
- Gibson, E. J. (1969). *Principles of perceptual learning and development*. NuevaYork: Appleton-Century Crofts.
- González-Peiteado, M. (2013). Los estilos de enseñanza y aprendizaje como soporte de la actividad docente. *Journal of Learning Styles*, 11, 11, 51-70.
- Grasha, A. F. (1996). *Teaching with Style. A practical guide to enhancing learning by understanding teaching and learning style*. Pittsburgh: Alliance Publishers.

- Gregorc, A. F. (1982). *Gregorc Style Delineator: development, technical and administration manual*. Columbia, CT: Gregorc Associates Inc.
- Herrmann, N. (1982). *Herrmann Brain Dominance Instrument*. Applied Services
- Herrmann, N. (1989). *The creative brain*. North Carolina: Brain Books, The Ned Herrmann Group.
- Hodgkinson, G. P. y Sadler-Smith, E. (2003). Reflections on reflections ... on the nature of intuition, analysis and the construct validity of the Cognitive Style Index. *Journal of Occupational and Organizational Psychology*, 76, 279-281.
- Honey, P. y Mumford, A. (1982). *Manual of Learning Styles*. London: P. Honey.
- Hosford C. C. y Siders W. A. (2010). Felder-Soloman's Index of Learning Styles: internal consistency, temporal stability and factor structure. *Teach Learn Med.* 22, 4, 298-303. doi: 10.1080/10401334.2010.512832.
- Huang, E. Y., Lin, S. W. y Huang, T. K. (2012). What type of learning style leads to online participation in the mixed-mode e-learning environment? A study of software usage instruction. *Computers y Education*, 58, 1, 338-349.
- Hughes, M., Hughes, P. y Hodgkinson, I. R. (2016). In pursuit of a 'whole-brain' approach to undergraduate teaching: implications of the Herrmann brain dominance model. *Studies in Higher Education*, 42, 12, 1-17. doi: 10.1080/03075079.2016.1152463
- Jackson, C. J. (2002). *Manual of the Learning Styles Profiler*. Sydney: Cymeon Research.
- Joy, S. y Kolb, D. A. (2009). Are there cultural differences in learning style? *International Journal of Intercultural Relations*, 33, 1, 69-85.
- Kagan, J. (1966). Reflection-impulsivity: the generality and dynamics of conceptual tempo. *Journal of Abnormal Psychology*, 71, 1, 17-24.
- Ka-J, W. y Teo, A. (2016). Communication Strategy Use in an Oral Narrative Task among English Learners with Different Hemispheric Brain Dominance. *LEARN Journal: Language Education and Acquisition Research Network*, 9, 2, 188-198.
- Kalantari, M., Tahan, M. y Taraghi, A. (2016). Study of Innovation in Learning Styles of Students in Different Secondary School Branches. *Journal of Psychology and Psychotherapy*, 6, 274. doi: 10.4172/2161-0487.1000274

- Kappe, F. R., Boekholt, L., Den Rooyen, C. y Van der Flier, H. (2009). A predictive validity study of the Learning Style Questionnaire (LSQ) using multiple, specific learning criteria. *Learning and Individual Differences*, 19, 4, 464-467.
- Kickul, J., Gundry, L. K., Barbosa, S. D. y Whitcanack, L. (2009). Intuition versus analysis? Testing differential models of cognitive style on entrepreneurial self-efficacy and the new venture creation process. *Entrepreneurship Theory and Practice*, 33, 2, 439-453. doi: 10.1111/j.1540-6520.2009.00298.x
- Knapp, T. R. (1998). Test review of Adaptive Style Inventory. En: J. C. Impara y B. S. Plake (Eds.), *The thirteenth mental measurements yearbook*. "Recuperado de: <http://marketplace.unl.edu/buros/>"
- Kolb, A. Y y Kolb, D. A. (2013). *The Kolb Learning Style Inventory, 4.0*. A comprehensive guide to the theory, psychometrics, research on validity and educational applications. Experience Based Learning Systems. "Recuperado de [https://www.researchgate.net/profile/David\\_Kolb/publication/303446688\\_The\\_Kolb\\_Learning\\_Style\\_Inventory\\_40\\_Guide\\_to\\_Theory\\_Psychometrics\\_Research\\_Applications/links/57437c4c08ae9f741b3a1a58/The-Kolb-Learning-Style-Inventory-40-Guide-to-Theory-Psychometrics-Research-Applications.pdf](https://www.researchgate.net/profile/David_Kolb/publication/303446688_The_Kolb_Learning_Style_Inventory_40_Guide_to_Theory_Psychometrics_Research_Applications/links/57437c4c08ae9f741b3a1a58/The-Kolb-Learning-Style-Inventory-40-Guide-to-Theory-Psychometrics-Research-Applications.pdf)". "
- Kolb, D. A. (1984). *Experimental learning: Experience as the source of learning and development*. Englewood Cliffs, NJ: Prentice-Hall.
- Kolb, D. A. (1999a) *Learning Style Inventory, Version 3*. Boston, MA: Training Resources Group. Kolb, D. A. (1999b) *Learning Style Inventory Version 3: Technical Specifications*. Boston, MA: Training Resources Group.
- Lau, W. W. F., Yuen, A. H. K. y Chan, A. (2015). *Variable-Centered and Person-Centered Approaches to Studying the VARK Learning Style Inventory*. Singapore: Springer,.
- Leite, W. L., Svinicki, M. y Shi, Y. (2010). Attempted Validation of the Scores of the VARK: Learning Styles Inventory with Multitrait-Multimethod Confirmatory Factor Analysis Models. *Educational and Psychological Measurement*, 70, 323-339.
- Manolis, C., Burns, D. J., Assudani, R. y Chinta, R. (2013). Assessing experiential learning styles: A methodological reconstruction and validation of the Kolb Learning Style Inventory. *Learning and individual differences*, 23, 44-52.

- Martínez, G. y Manzo, S. (2013). *Aplicación del modelo cuadrante cerebral de Herrmann y su relación con los estilos de aprendizaje*. XIV Congreso virtual de psiquiatría.com Interpsiquis, 1-28 febrero.
- Martínez-Fernández, J. R., García-Ravidá, L., González-Velázquez, L., Gutiérrez-Braojos, C., Poggioli, L., Ramírez- Otálvaro, P. y Tellería, M. B. (2009). *Inventario de Estilos de Aprendizaje en español*. Documento interno del Grup de Recerca PAFIU. Barcelona: Universitat Autònoma de Barcelona. “Recuperado de <http://grupsderecerca.uab.cat/pafiu/>”
- Meneely, J. (2010). Educating adaptable minds: how diversified are the thinking preferences of interior design students? *Journal of Interior Design*, 35, 3, 21-32. doi: 10.1111/j.1939-1668.2010.01040.x
- Miller, M., De Clerck, J. P., Sorby, S. A., Roberts, L. M., Endres, W. J. y Hale, K. D. (2013). Meeting the NAE Grand Challenge: Personalized Learning for Engineering Students through Instruction on Megacognition and Motivation Strategies. *120th ASEE Annual Conference Proceedings*, Atlanta, Georgia, USA. 23-26 junio.
- Miller, M., Prabhakara, P. y Sorby, S. A. (2015). Reflection and Evaluation Data from e-Learning Modules on Learning Styles and Motivation. *Proc. ASEE Annual Conference and Exposition*, Seattle, WA, USA, 14-17 junio.
- Miller, M., Sorby, S. y De Clerck, J. (2015). E-Learning Modules for Improving Lifelong Learning Ability, *Proc. ASEE Annual Conference and Exposition*, Seattle, WA, USA, 14-17 junio.
- Mortensen, C. J., Thoron, A. C. y Miot, J. K. (2015). Current learning styles of undergraduate animal-studies students in a 2-year and 4-year degree program. *NACTA Journal*, 59, 1, 75-80.
- Mutua, M. N. (2015). *A correlation study between learning styles and academic achievement among secondary school students in Kenya* (Tesis inédita de doctorado). Universidad de Nairobi. Kenia.
- Myers, I. B. y McCaulley, M. H. (1985). *Manual: a guide to the development and use of the Myers-Briggs Type Indicator*. Palo Alto, CA: Consulting Psychologists Press.
- Myers, I. B. y McCaulley, M. H. (1998). *MTBI manual: a guide to the development and use of the Myers-Briggs Type Indicator*. Palo Alto, CA: Consulting Psychologists Press.
- Njoku, J. N. y Abdulhamid, B. (2016). Preference of Learning Styles and its Relationship with Academic Performance among Junior Secondary School Students in Dutse Local Government Area, Jigawa State,

- Nigeria. *International Journal of Education and Practice*, 4, 3, 127-133.
- O'Connor, P. J., y Jackson, C. J. (2008). The factor structure and validity of the Learning Styles Profiler (LSP). *European Journal of Psychological Assessment*, 24, 2, 117-123. doi: 10.1027/1015-5759.24.2.117
- O'Mahony, S. M., Sbayeh, A., Horgan, M., O'Flynn, S. y O'Tuathaigh, C. M. (2016). Association between learning style preferences and anatomy assessment outcomes in graduate-entry and undergraduate medical students. *Anatomical sciences education*, 9, 4, 391-399. doi: 10.1002/ase.1600
- Ovington, L. A., Saliba, A. J. y Goldring, J. (2016). Dispositional Insight Scale: development and validation of a tool that measures propensity toward insight in problem solving. *Creativity Research Journal*, 28, 3, 342-347.
- Parker, J. A. (1985). Test review of Barsch Learning Style Inventory. En: J. V. Mitchell (Ed.), *The ninth mental measurements yearbook*. "Recuperado de <http://marketplace.unl.edu/buros/>"
- Pashler, H., McDaniel, M., Rohrer, D. y Bjork, R. (2008). Learning styles: Concepts and evidence. *Psychological Science in the Public Interest*, 9, 3, 103-119.
- Penger, S., Tekavcic, M. y Dimovski, V. (2011). Comparison, validation and implications of learning style theories in higher education in Slovenia: an experiential and theoretical case. *International Business y Economics Research Journal*, 7, 12, 25-44.
- Peterson, K., DeCato, L. y Kolb, D. A. (2015). Moving and learning: Expanding style and increasing flexibility. *Journal of Experiential Education*, 38, 3, 228-244.
- Prien, E. P. (1998). Test review of Adaptive Style Inventory. En J. C. Impara y B. S. Plake (Eds.), *The thirteenth mental measurements yearbook*. "Recuperado de <http://marketplace.unl.edu/buros/>"
- Pupo, E. A. (2012). Los estilos de enseñanza, una necesidad para la atención de los estilos de aprendizaje en la educación universitaria. *Journal of Learning Styles*, 10, 10, 79-87.
- Quiñones, C. (2004). Metodología de Estrategia Enseñanza-Aprendizaje y Estilos de Aprendizaje. *UMBRAL. Revista de Educación, Cultura y Sociedad*, 6, 48-61.

- Reio, T. G. y Wiswell, A. K. (2006). An examination of the factor structure and construct validity of the Gregorc Style Delineator. *Educational and Psychological Measurement*, 66, 3, 489-501.
- Reynolds, C. R. (2003). Test review of Learning Style Inventory, Version 3. En B. S. Plake, J. C. Impara, y R. A. Spies (Eds.), *The fifteenth mental measurements yearbook*. “Recuperado de <http://marketplace.unl.edu/buros/>”
- Riding, R. (1991). *Cognitive Styles Analysis – CSA administration*. Birmingham: Learning and Training Technology.
- Rodríguez, A. L., Tijerina, B. A. T. y García, J. L. G. (2016). Implementación del instrumento Quirontest para medir estilos de aprendizaje en estudiantes de pregrado en línea. *Journal of Learning Styles*. 9, 17, 240-267.
- Saarikoski, L., Salojärvi, S., Del Corso, D, y Ovcin, E. (2001). *The 3DE: An Environment for the Development of Learner-Oriented Customised Educational Packages*. International Conference on Information Technology Based Higher Education and Training (ITHET): Kumamoto, 4-6 julio.
- Sadler-Smith, E., Spicer, D. P. y Tsang, F. (2000). Validity of the Cognitive Style Index: replication and extension. *British Journal of Management*, 11, 2, 175-181. doi: 10.1111/1467-8551.t01-1-00159
- Schar, M. (2015). *Pivot thinking: Predicting entrepreneurial intent among engineering students and faculty using problem solving style preference*. Frontiers in Education Conference (FIE). El Paso, Texas, USA, 21-24 octubre.
- Shaikh, V. y Waychal, P. (2015). Experience of using felder–soloman index of learning styles. *Proceedings of the International Conference on Transformations in Engineering Education*, 369-378. doi: 10.1007/978-81-322-1931-6\_43
- Shinnick, M. A. y Woo, M. A. (2015). Learning style impact on knowledge gains in human patient simulation. *Nurse Education Today*, 35, 1, 63-67. doi: 10.1016/j.nedt.2014.05.013
- Shukr, I., Zainab, R. y Rana, M. H. (2013). Learning styles of postgraduate and undergraduate medical students. *Journal of College of Physicians and Surgeon Pakistan*, 23, 1, 25-30. doi: 01.2013/JCPSP.2530.
- Shum, D. (2003). Test review of Learning Style Inventory, Version 3. En: B. S. Plake, J. C. Impara, y R. A. Spies (Eds.), *The fifteenth mental measurements yearbook*. “Recuperado de <http://marketplace.unl.edu/buros/>”

- Singh, L., Govil, P. y Rani, R. (2015). Learning Style preferences among secondary school students. *International Journal of Recent Scientific Research*, 6, 5, 3924-3928.
- Sternberg, R.J. (1999). *Thinking styles*. Cambridge: Cambridge University Press.
- Tracz, W. (2017). *An exploratory study of Ontario college electrical engineering students learning styles*. (Tesis inédita de doctorado). Universidad de Athabasca. Canadá.
- van Heerden, H., van Eck, E., Burger, M. y Khan, R. (2016). Learning Style Preferences of Undergraduate Quantity Surveying Students: Quantity Surveying. *HASSACC-Human and Social Sciences at the Common Conference*. 3-7 octubre.
- Velásquez, A. M. V., Ortiz, J. F. Z. y Rodríguez, A. L. (2016). La relación entre los estilos de aprendizaje y el rendimiento académico en matemáticas en alumnos de ciclo v de educación secundaria. *Journal of Learning Styles*, 9, 18, 70-94.
- Vermunt, J. D. (1996). Metacognitive, cognitive and affective aspects of learning styles and strategies: a phenomenographic analysis. *Higher Education*, 31, 1, 25-50.
- Villalobos, L., González, M., Muñoz, J., Ostoic, C. y Veliz, N. (2016). *Estilos de aprendizaje de Honey-Alonso en estudiantes de ingreso universidad Antofagasta 2012-2013-2014-2015*. V Congreso CLABES. Talca, Chile, 11-13 noviembre.
- Wang, J. y Mendori, T. (2015). The Reliability and Validity of Felder-Silverman Index of Learning Styles in Mandarin Version. *Information Engineering Express International Institute of Applied Informatics*, 1, 3, 1-8.
- Wilkinson, T., Boohan, M. y Stevenson, M. (2014). Does learning style influence academic performance in different forms of assessment?. *Journal of anatomy*, 224, 3, 304-308. doi: 10.1111/joa.12126.
- Wismath, S., Orr, D. y Good, B. (2014). Metacognition: Student reflections on problem solving. *Journal on Excellence in College Teaching*, 25, 2, 69-90.
- Wismath, S., Orr, D. y Zhong, M. (2014). Student perception of problem solving skills. *Transformative Dialogues*, 7, 3, 1-17.

Wium, A. M., Pitout, H., Human, A. y du Toit, P. H. (2015). An analysis of thinking preferences across three health care disciplines. *Innovations in Education and Teaching International*, 1-9.

Zywno, M. (2003). *A Contribution to Validation of Score Meaning for Felder Soloman's Index of Learning Styles. Annual Conference, Nashville, Tennessee, 22-25 junio.*

**Información de contacto:** M<sup>a</sup> Luz Diago Egaña. Universidad Internacional de la Rioja. Facultad de Educación. Departamento de Didáctica de las Matemáticas y de las Ciencias Experimentales. Avenida de la Paz, 137. 26006, Logroño. (La Rioja). España. E-mail: marialuz.diago@unir.net

## Anexo

**TABLA I.** Palabras clave, instrumentos y acrónimos empleados en las búsquedas en las distintas bases de datos. Todos los términos se manejaron en dos idiomas (castellano e inglés).

<b>PALABRAS CLAVE</b>	<b>INSTRUMENTOS</b>	<b>ACRÓNIMOS</b>
Desempeño académico Hemisferios cerebrales y aprendizaje Preferencias de aprendizaje Estrategias de aprendizaje Estilos de aprendizaje Estilos de aprendizaje y mitos Neurociencia Canales de percepción Enfoques y habilidades de estudio Estilos cognitivos Cuestionario Instrumento de medición Estilos de enseñanza	Inventario para estudiantes Inventario Barsch estilos de aprendizaje Índice de estilo cognitivo Cuestionario de Evaluación de Procesos de Estudio y Aprendizaje Cuestionario Escanero-Soria de estilos de aprendizaje Cuestionario Honey-Alonso de Estilos de Aprendizaje Estilo de uso del Espacio Virtual Modelo de estilos mentales de Gregorc y Delinador de estilo Instrumento de dominio del cerebro de Herrmann Índice de estilos de aprendizaje Inventario de Estilos de aprendizaje Inventario de Estilos adaptativos de Kolb Inventario de Estilos de aprendizaje de Kolb Inventario de estilos de aprendizaje Perfil de Estilos de aprendizaje Cuestionario de Estilos de Aprendizaje Escala de estilos de Aprendizaje Perfil de Estilo Motivacional Indicador Myers-Briggs Test Quiron Cuestionario del proceso de estudio Test de Barsch	ASI ASSIST BLSI CEPEA CESEA CHAEA CSI GSD HBDI ILS KLSI LSI LSP LSQ LSS MBTI MSP PNL SPQ VAK VARK

Fuente: elaboración propia

**TABLA II.** Características del cuestionario CSI utilizado para medir “Estilos de Aprendizaje”, que según Coffield et ál. (2004) reúne 4 criterios mínimos.

<b>Nombre del cuestionario</b>	<b>Referencia</b>	<b>Constructo</b>	<b>Dimensiones</b>	<b>Validado (Sí o No)</b>	<b>Características psicométricas</b>
CSI (Cognitive Style Index)	Allinson y Hayes (1996).	Preferencias para procesar la información	Estilo intuitivo, Estilo analítico	Sí (Sadler-Smith et ál. 2000; Coffield et ál. 2004; Armstrong y Qi, 2016).	consistencia interna, fiabilidad test-retest, validez de constructo y validez predictiva

Fuente: elaboración propia

**TABLA III.** Características de los cuestionarios utilizados para medir “Estilos de Aprendizaje” que según Coffield et ál. (2004) reúnen 3 y 2 criterios mínimos.

<b>Nombre del cuestionario</b>	<b>Referencia</b>	<b>Constructo</b>	<b>Dimensiones</b>	<b>Validado (Si o No)</b>	<b>Características psicométricas</b>
MSP (Motivational Style Profile)	Apter, Mal-lows y Williams (1998)	Características de la personalidad (5 pares de estados motivacio-nales y tendencias hacia la excitación, el esfuer-zo y el optimismo / pesimismo)	télico/paratélico – negativismo/confor-midad – maestría/simpatía auténticas maestría/simpatía aloicas Evitar/Buscar la excitación Optimismo/pesi-mismo Excitación /Esfuerzo	SI (Coffield et al, 2004)	consisten-cia interna, fiabilidad test-retest y validez predictiva
ILS (Inventory of Learning Styles)	Vermunt (1996)	Estrategias y aproxima-ciones de Aprendizaje	Dirigido al signifi-cado, dirigido a la aplicación, dirigido a la reproducción y no dirigido	SI (Coffield et ál. 2004)	consisten-cia interna, fiabilidad test-retest y validez del constructo (Coffield et ál. 2004)
ASSIST (Ap-proaches and Study Skills Inventory for Students)	Entwis-tle (1988)	Estrategias y aproxima-ciones de Aprendizaje	Aproximación profunda, superficial y estratégica	SI (Coffield et ál. 2004)	consistencia interna, y validez del constructo
HBDI (Her-rmann’s Brain Dominance Instrument)	Her-rmann (1989)	Estilos de pensamiento/ Preferencias apren-dizaje	Pensamiento ana-lítico, secuencial, interpersonal e imaginativo	SI (Cof-field et ál. 2004); Bunder-son, 1994)	fiabilidad test-retest y validez del constructo (Coffield et ál., 2004)
MBTI (Myers-Briggs Type Indicator)	Myers y Mc-Caulley (1985). Myers y Mc-Caulley (1998).	Características de la personalidad, Estructu-ra cognitiva	Percibir/juzgar Sensa-ción/intuición, pensa-miento /sentimiento y Extraversión/intro-versión	SI (Coffield et ál. 2004)	consistencia interna y fiabilidad test-retest

Fuente: elaboración propia

**TABLA IV.** Características de los cuestionarios utilizados para medir “Estilos de Aprendizaje”, que según Coffield et ál. (2004) reúnen 1 criterio mínimo.

Nombre del cuestionario	Referencia	Constructo	Dimensiones	Validado (Sí o No)	Características psicométricas
LSQ (Learning style questionnaire)/ CHAEA (adaptación al castellano)	Honey y Mumford (1992)	Preferencias de Estilos de aprendizaje	Activo, reflexivo, teórico y pragmáticos	SI (Coffield, et ál. 2004.); NO (Kappe, Boekholt, Den Rooyen, y Van der Flier, 2009).	confiabilidad test-retest confiabilidad test-retest
VAK/ VARK/PNL	Dunn y Dunn (1978); Dunn, Dunn, y Price, (1989)/ Fleming, (2001)	Modalidades en la percepción de la información	Visual, auditivo y kinestésico/Visual, auditivo, lector y kinestésico	SI (Coffield et ál. 2004); Leite, Svinicki y Shi (2010).	Coffield et al. (2004): validez predictiva.  Leite et ál. (2010): fiabilidad. Advierten de tener precaución de su utilización en investigación en lo referente a la redacción de los ítems y la puntuación de los algoritmos de la escala
LSI Version 3	Kolb (1999a y b)	Preferencias de EA Experimental	Acomodador, Divergente, Asimilador, Convergente	SI (Coffield et ál. 2004) NO (Reynolds, 2003; Shum, 2003).	confiabilidad test-retest -----
GSD (Gregorc's Mind Styles Model and Style Delineator)	Gregorc (1982)	Canales específicos para recibir, procesar y expresar información	Concreto secuencial/ abstracto, Aleatorio- resumen Secuencial/ concreto al azar	SI (Coffield et ál. 2004) NO (Reio y Wiswell, 2006)	Validez predictiva Coeficiente alfa de Cronbach
LSP (Jackson's Learning Styles Profiler)	Jackson (2002)	Características estables de la personalidad	Iniciador/Razonador/Analista/Implementador	SI (O'Connor, y Jackson, 2008)	Factor estructura

Fuente: elaboración propia

**TABLA V.** Herramientas utilizadas para medir “Estilos de Aprendizaje” no seleccionados por Coffield et ál. (2004) con sus características métricas.

Nombre del cuestionario	Referencia	Constructo	Dimensiones	Validado (Sí o No)	Características psicométricas
SPQ (Study-Process Questionnaire) CEPEA (Cuestionario de Evaluación de Procesos de Estudio y Aprendizaje)	Biggs (1987) Barca (1999)	Motivación y estrategias de aprendizaje	Superficial, profundo y de logro	SI (Burnett y Dart, 2000)	Validez y confiabilidad
Test de Barsch (Barsch Learning Style Inventory)	Barsch, (1996).	Preferencias de aprendizaje	Visual, Auditivo, Táctil	NO (Biggs, 1985; Parker, 1985).	-----
KLSI 4 (Kolb Learning Styles Inventory, Version 4)	Kolb y Kolb (2013)	Mide el grado en que una persona cambia su Estilo de Aprendizaje Experimental según el contexto de aprendizaje	Iniciando, Experimentando, Imaginando, Reflejando, Analizando, Pensando, Decidiendo, Actuando y Balanceando	SI (LSI Version 4: Experience Based Learning Styles, Inc)	Confiabilidad, validez interna y externa
ASI (Kolb Adaptive Style Inventory)	<a href="http://learningfromexperience.com/tools/kolb-adaptive-style-inventory-asi/">http://learningfromexperience.com/tools/kolb-adaptive-style-inventory-asi/</a>	Mide flexibilidad adaptativa (cómo adaptamos el EA a 4 contextos diferentes de aprendizaje)	Actuar. Valorar, Pensar; Decidir	NO (Knapp, 1998; Prien, 1998)	-----

ILS (Index of Learning Styles)	Felder y Soloman, (1997)	Preferencias o tendencias de aprendizaje	Sensorial o intuitivo Visual o verbal Activo o reflexivo Secuencial o global	SI (Zywno, 2003;  Felder y Spurlin, 2005;  Hosford y Siders 2010;  Wang y Mendori, 2015;  Al-Azawei, Parslow, y Lundqvist, 2015) NO (Çardak, y Selvi, 2016)	Confiabilidad test-retest, factor estructura, validez interna, correlación total y correlación interescala Confiabilidad y validez  Consistencia interna, estabilidad temporal y factor estructura  Validez interna, correlación interescala y validez del constructo Confiabilidad y validez de percepción moderadas -----
CESEA (Cuestionario Escanero-Soria de estilos de aprendizaje)	Escanero-Marcén, Soria, Guerra-Sánchez y Silva (2016)	Estilos de aprendizaje (estilos cognitivos)	Teórico (conocer), Dinámico (probar), Reflexivo (pensar), Operativo (hacer)	NO	-----
Estilo de uso del Espacio Virtual	Barros (2010)	Estilo de uso del espacio virtual	Participativo, Buscador e Investigador, Estructurador y planeador, Concreto y Productivo	NO	-----
Quiron Test	Rodríguez, Tijerina y García (2016)	Preferencias a la hora de cursar asignaturas on line o en modalidad híbrida on line-presencial	Analítico/Global; Dependiente/independiente; Teórico/práctico; Visual/verbal	SI	Alfa de Cronbach
LSS (Learning Styles Scale)	Abdollahimohammad y Ja'afar (2014)	Preferencias de aprendizaje	Perceptivo, solitario, analítico, competitivo, e imaginativo	SI (Abdollahimohammad y Ja'afar, 2014)	Coefficiente alfa de Cronbach, test-retest

Fuente: elaboración propia

**TABLA VI.** Utilización en los últimos años de los distintos cuestionarios.

<b>Nombre del cuestionario</b>	<b>Algunos Trabajos posteriores a 2004 que lo han empleado</b>
CSI	(Barbosa, Gerhardt y Kickul, 2007; Kickul, Gundry, Barbosa y Whitcanack, 2009; Ovington, Saliba y Goldring, 2016; Armstrong y Qi, 2016)
HBDI	(Meneely, 2010; De Boer, Bothma y du Toit, 2011; Martínez y Manzo, 2013; Bothma y de Boer, 2015; Schar, 2015; Wium, Pitout, Human, y du Toit, 2015; Hughes, Hughes y Hodgkinson, 2016; Ka-j y Teo, 2016; van Heerden, van Eck, Burger y Khan, 2016)
LSQ	(Penger, Tekavcic y Dimovski, 2011; Gargallo, Almerich, Suárez, García y Garfella, 2012; Shukr, Zainab, y Rana, 2013; Wilkinson, Boohan y Stevenson, 2014; Aljaberi, 2015; O'Mahony, Sbayeh, Horgan, O'Flynn y O'Tuathaigh, 2016; Villalobos, González, Muñoz, Ostoic y Veliz, 2016).
VAK/VARK	(Penger, Tekavcic y Dimovski, 2011; Al Hamdani, 2015; Bhagat, Vyas y Singh, 2015; Tracz, 2017)/(Díaz y Aly, 2015; Lau, Yuen y Chan, 2015; Velásquez, Ortiz, y Rodríguez, 2016)
Gergorc Style Delineator (GSD)	(Mortensen, Thoron y Miot, 2015)
Test de Barsch	(Miller, De Clerck, Sorby, Roberts, Endres y Hale, 2013; Wismath, Orr y Good, 2014; Wismath, Orr y Zhong, 2014; Miller, Prabhakara y Sorby, 2015; Miller, Sorby y De Clerck, 2015; Mutua, 2015; Singh, Govil y Rani, 2015; Alavi y Makarem, 2016; Njoku y Abdulhamid, 2016)
ILS (Felder-Soloman)	(Huang, Lin y Huang, 2012; Cheng, 2014; Shaikh y Waychal, 2015)
CESEA	(Escanero-Marcén, Soria, Guerra-Sánchez y Silva, 2016)
Reduced Kolb Learning Style Inventory (RLSI)	(Manolis, Burns, Assudani y Chinta, 2013)
KLSI	(Peterson, DeCato y Kolb, 2015; Kalantari, Tahan y Taraghi, 2016)
LSI	(Donche, Coertjens y Van Petegem, 2010; Shinnick y Woo, 2015)
LSS	(Abdollahimohammad y Ja'afar, 2014; Abdollahimohammad y Ja'afar, 2015)

Fuente: elaboración propia

TABLA VII. Resumen de los instrumentos seleccionados para medir "Estilos de Aprendizaje"

Nombre del cuestionario/ Referencia	Constru- cto	Dimen- siones	Descripción de las di- mensiones	Validado (Si o No)	Características psicométricas
ILS (Index of Learning Styles)/ Felder y Soloman, (1997)	Prefe- rencias o tenden- cias de aprendi- zaje	Sensiti- vo/Intui- tivo	Sensitivo: Concreto, práctico, orientado hacia hechos y procedimien- tos	SI (Zywno, 2003);  Felder y Spurlin, 2005;  Hosford y Siders 2010;  Wang y Mendori, 2015;  Al-Azawei, Pars- low, y Lundqvist, 2015) NO (Çardak, y Selvi, 2016)	Confiabilidad test-retest, factor estructura, validez interna, correla- ción total y corre- lación interescala  Confiabilidad y validez  Consistencia interna, estabilidad temporal y factor estructura Validez interna, correlación inte- rescala y validez del constructo  Confiabilidad y validez de percep- ción moderadas  -----
			Intuitivo: Conceptual, innovador, orientad hacia las teorías y significados subyacentes		
		Visual/ Verbal	Visual: Prefieren representa- ciones visuales de material presentado - imágenes, diagramas, diagramas de flujo...-		

			Verbal: Prefieren explicaciones escritas y orales		
		Activo/ Reflexivo	Activo: Aprenden experimentando, disfrutan trabajando con otros		
			Reflexivo: Aprenden pensando las cosas, prefieren trabajar solos o con una o dos personas		
		Secuencial/ Global	Secuencial: Lineal, ordenados, aprenden en pequeños pasos incrementales		
			Global: Holísticos, pensadores sistémicos, aprenden en grandes saltos		
VAK/VARK/PNL/ Dunn y Dunn (1978); Dunn, Dunn, y Price, (1989)/Fleming, (2001)	Modalidades en la percepción de la información	Visual	Prefieren ver la información presentada de forma visual en lugar de escrita, por ejemplo a través de imágenes, cuadros, diagramas, círculos, flechas, láminas...	SI (Coffield et ál. 2004; Leite, Svinicki y Shi (2010).	Coffield et al, (2004): validez predictiva.  Leite et ál. (2010): fiabilidad. Advierten de tener precaución de su utilización en investigación en lo referente a la redacción de los ítems y la puntuación de los algoritmos de la escala

		Auditivo	Aprenden mejor escuchando la información (en exposiciones orales, conferencias, debates...). Recuerdan con facilidad lo que escuchan y lo que expresan verbalmente.		
		Lecto/ Escritor	Prefieren todo lo que se relacione con leer y escribir y recabar la información que se muestra como palabras. Los materiales de aprendizaje que están principalmente basados en texto son muy preferidos por estos estudiantes.		
		Kinestésico	Aprenden mejor tocando y haciendo. La experiencia práctica es importante para ellos.		

Fuente: elaboración propia



# Procesamiento de magnitudes numéricas y ejecución matemática<sup>1</sup>

## Numerical magnitude processing and mathematical achievement

DOI: 10.4438/1988-592X-RE-2017-381-383

Jose txu Orrantia  
Sara San Romualdo  
Rosario Sánchez  
Laura Matilla

*Universidad de Salamanca*

David Muñez

*National Institute of Education, Center for Research in Child Development (Singapur)*

Lieven Verschaffel

*Katholieke Universiteit Leuven, Center for Instructional Psychology and Technology*

### Resumen

Recientes investigaciones sugieren que las diferencias individuales en ejecución matemática están relacionadas con las habilidades de procesamiento numérico básico, tales como la capacidad para procesar magnitudes numéricas. Una cuestión clave en este reciente campo de investigación es qué habilidades relacionadas con el procesamiento de magnitudes predicen la ejecución en matemáticas, el procesamiento de magnitudes no simbólicas o el acceso a esas magnitudes desde los números simbólicos. En este estudio extendemos esta investigación analizando el rol del tamaño de las magnitudes utilizando un diseño predictivo longitudinal. Cincuenta y dos participantes de 1º de

---

<sup>(1)</sup> Este trabajo fue realizado como parte del proyecto financiado PSI2015-66802-P del Ministerio de Economía y Competitividad

Educación Primaria fueron evaluados en tareas de procesamiento de magnitudes numéricas, tanto simbólicas como no simbólicas con cantidades grandes y pequeñas, y dos años después se les evaluó en ejecución matemática. Los análisis de regresión jerárquica muestran que el procesamiento de magnitudes simbólicas de cantidades grandes (dos dígitos) es un predictor más robusto de la futura ejecución matemática que las demás medidas de procesamiento de magnitudes. Estos resultados se interpretan en términos de sus implicaciones educativas, específicamente en aspectos relacionados con la identificación temprana de estudiantes en riesgo de presentar dificultades en el aprendizaje de las matemáticas, algo prioritario en cualquier sistema educativo desde el punto de vista de la prevención.

*Palabras clave:* ejecución matemática, educación primaria, procesamiento numérico, procesamiento de magnitudes numéricas, sistema numérico aproximado

### **Abstract**

Recent research suggests that individual differences in mathematics are related to the ability to basic number processing skills, such as the ability to process numerical magnitudes. A key question in this emerging field of research is which skills related to the magnitude processing predict the mathematical competence: either no symbolic magnitude processing, or the access to those magnitudes from the symbolic numbers. The present study extended this research by investigating the role of the size of the quantities (small vs. large). Fifty-two children were assessed on nonsymbolic and symbolic magnitude processing measures at the start of formal schooling and mathematics achievement was evaluated two years later. Hierarchical regression analyzes showed that large symbolic magnitude processing was a stronger predictor of future mathematical achievement compared to the other magnitude processing measures. These results were interpreted in terms of their educational implications, specifically in the use of screening tools for identifying children with difficulties in mathematics.

*Keywords:* mathematical achievement, elementary school, numerical processing, numerical magnitude processing, approximate number system

## **Introducción**

El estudio de las diferencias individuales en matemáticas ha producido un importante número de investigaciones centradas en analizar qué factores subyacen a dichas diferencias, factores que van desde habilidades

cognitivas de dominio general, relacionadas, entre otras, con la memoria de trabajo, las funciones ejecutivas o la velocidad de procesamiento, hasta habilidades específicas del dominio matemático (LeFevre, Wells, y Sowinski, 2016). Entre estas últimas, el procesamiento de magnitudes numéricas ha generado un considerable número de estudios en el último lustro (véase Lyons y Ansari, 2015, para una revisión), planteándose la hipótesis de que las habilidades para procesar y representar magnitudes numéricas se relacionan con la ejecución en matemáticas. El presente artículo pretende analizar dicha cuestión. Concretamente, el objetivo de este trabajo es analizar, en un estudio longitudinal, hasta qué punto las diferencias individuales en el procesamiento y representación de magnitudes numéricas, en niños que están comenzando con el aprendizaje formal de la aritmética, predice la ejecución aritmética dos años después, cuando se han consolidado las habilidades aritméticas básicas. En lo que sigue se revisan las investigaciones más recientes relacionados con esta cuestión, para, desde las limitaciones encontradas en las mismas, presentar el actual estudio.

## Procesamiento de la magnitud numérica y ejecución matemática

Hay evidencias de que el ser humano (al igual que otras especies animales) está dotado de un mecanismo que le permiten representar y procesar magnitudes numéricas no simbólicas desde que nace (Dehaene, 2011). Este mecanismo es conocido como sistema numérico aproximado (SNA en adelante), un sistema primitivo de representación no verbal que permite a los individuos discriminar magnitudes de manera aproximada y cuyo desarrollo no depende de una enseñanza explícita (Feigenson, Dehaene, y Spelke, 2004). La ejecución en discriminación entre magnitudes se rige por la ley de *Weber*; según la cual la discriminación entre dos magnitudes depende de su ratio más que de su diferencia absoluta. La tarea más ampliamente utilizada para analizar la precisión del SNA ha sido la comparación de magnitudes no simbólicas (e.g., Libertu, Feigenson y Halberda, 2011), en la que se presenta a los participantes dos conjuntos de puntos para decidir dónde hay más. En esta tarea, el tiempo de respuesta y los errores disminuyen cuando la ratio o la distancia entre las magnitudes aumenta. Se asume que las magnitudes se representan de una manera aproximada (i.e., como una distribución Gaussiana

alrededor de una magnitud específica) a lo largo de una hipotética línea numérica mental, y que los efectos de ratio y distancia se producen por un mayor solapamiento representacional de las magnitudes más cercanas. En este sentido, el tamaño de los efectos de distancia y ratio se consideran indicadores de la precisión de las representaciones de la magnitud numérica, y dicho tamaño disminuye con la edad, llevando a representaciones cada vez más precisas (Halberda y Feigenson, 2008).

A pesar de que el SNA se diferencia de los sistemas numéricos simbólicos que permiten representar cantidades de manera precisa, se ha sugerido que el SNA es la base sobre la que se asientan las habilidades numéricas simbólicas (Dehaene, 2011). Se ha propuesto que cuando se aprenden los símbolos para representar números, éstos adquieren significado cuando se proyectan en el SNA preexistente (e.g., Barth, Starr y Sullivan, 2009; Lipton y Spelke, 2005; Mundy y Gilmore, 2009; pero ver Reynvoet y Sasanguie, 2016, para una explicación alternativa). De hecho, los efectos de distancia y ratio también se producen en tareas de comparación de magnitudes simbólicas (en este caso la tarea es comparar dígitos arábigos para decidir cuál es mayor), lo que sugeriría también un solapamiento representacional en la línea numérica mental cuando se accede a la magnitud desde los números simbólicos. Aunque no está claro cómo se lleva a cabo la proyección de los símbolos en el SNA (ver Leibovich y Ansari, 2016, para una revisión de esta cuestión), se ha sugerido que el SNA juega un importante rol en el aprendizaje de las matemáticas simbólicas, y que por lo tanto las diferencias individuales en la ejecución del SNA se relacionan con las diferencias individuales en la ejecución matemática, aunque los resultados no son concluyentes (Lyons y Ansari, 2015).

Así, dos recientes meta-análisis (Chen y Li, 2014; Fazio, Bailey, Thompson, y Siegler, 2014) han mostrado una relación moderada, aunque significativa, entre el procesamiento de la magnitud numérica no simbólica y la ejecución en aritmética. Sin embargo, cuando se analizan conjuntamente el procesamiento de magnitudes numéricas no simbólicas con magnitudes numéricas simbólicas, los resultados son menos concluyentes. Por ejemplo, De Smedt, Noël, Gilmore y Ansari (2013) llevaron a cabo una revisión narrativa mostrando que solo el 44% de los estudios empíricos analizados en dicha revisión presentaron una relación significativa entre ejecución en aritmética y procesamiento de la magnitud numérica no simbólica, mientras que el 76% de los estudios

relevaron una asociación significativa con el procesamiento de la magnitud numérica simbólica. De la misma manera, otro reciente meta-análisis de Schneider et al. (en prensa) que incluyó tanto medidas de comparación de magnitudes numéricas no simbólicas como simbólicas reveló que el tamaño del efecto fue significativamente más grande para la asociación entre ejecución aritmética y comparación de magnitudes simbólicas.

Por lo tanto, los estudios han convergido en que el procesamiento de la magnitud numérica puede ser la base sobre la que se asienta el aprendizaje de las matemáticas. Sin embargo, los resultados de estos estudios no son concluyentes acerca de si es el procesamiento de magnitudes numéricas no simbólicas (e.g., puntos), el procesamiento de magnitudes numéricas simbólicas, (e.g., dígitos) o ambos, lo que se asocia con la ejecución en matemáticas (Orrantía et al., 2017). Una limitación de los estudios previos es que, en la mayoría de ellos, las tareas de comparación de magnitudes simbólicas incluyeron números entre 1 y 9 (Xenidou-Dervou, Molenaar, Ansari, van der Schoot, y van Lieshout, 2016). Sin embargo, el desarrollo del conocimiento de la magnitud numérica implica representaciones cada vez más precisas de un rango cada vez más amplio de números sobre la hipotética línea numérica mental (véase la Teoría Integrada del desarrollo numérico de Siegler, 2016). En este sentido, niños con representaciones relativamente precisas de números pequeños (por ejemplo, de un dígito), pueden presentar representaciones menos precisas de números más grandes (por ejemplo, de dos dígitos). Esto ha sido demostrado tanto en tareas de comparación de magnitudes simbólicas (e.g., Ashkenazi, Mark-Zigdon, y Henik, 2009), como en tareas de estimación en la línea numérica. En el caso de la recta numérica, el nivel de ejecución refleja la habilidad del participante para trasladar números simbólicos a magnitudes no simbólicas. Así, los participantes tienen que situar un número en una línea con 0 en un extremo y 10 o 100 en el otro (e.g., Bertelletti, Lucangeli, Piazza, Dehaene, y Zorzi, 2010). Por lo tanto, es importante analizar hasta qué punto el procesamiento de magnitudes numéricas se relaciona con la ejecución aritmética en tareas que incluyan números de dos dígitos, especialmente al inicio de la escolaridad formal, periodo educativo en el que se adquieren estos números.

El único estudio hasta la fecha que ha analizado esta cuestión con participantes que están comenzando con el aprendizaje formal de las

matemáticas fue el llevado a cabo por Xenidou-Dervou et al. (2016)<sup>2</sup>. Estos autores desarrollaron un estudio longitudinal en el que utilizaron tareas de comparación de magnitudes numéricas simbólicas y no simbólicas con cantidades pequeñas y grandes (1–9 vs. 6–70). Aplicaron estas pruebas a participantes en tres años sucesivos: Educación Infantil (*Kindergarten*), primero y segundo curso de Educación Primaria. Además, evaluaron la ejecución en matemáticas al final del segundo curso. Sus resultados mostraron que las distintas medidas utilizadas siguieron distintas trayectorias evolutivas, pero lo que es más importante, la comparación de magnitudes simbólicas con cantidades grandes fue el predictor más robusto de la ejecución en matemáticas, incluso por encima de la inteligencia. No obstante, las medidas utilizadas en las tareas de comparación de magnitudes fueron porcentaje de aciertos y tiempo de reacción (TR en adelante), que en el caso de las tareas simbólicas pueden reflejar el acceso al SNA desde los números simbólicos o, alternativamente, la eficacia para procesar información numérica simbólica *per se* (De Smedt et al., 2013). En este sentido, sería importante analizar hasta qué punto otras medidas más relacionadas con el acceso a la representación de la magnitud, como el efecto de distancia, son predictores de la ejecución en matemáticas. Además, la combinación de distintos tipos de medidas en un mismo estudio permitiría analizar cuánto de la varianza de la ejecución en matemáticas se atribuye a las diferencias individuales en la representación de la magnitud subyacente y cuánto a la eficacia para procesar magnitudes.

En este contexto, el objetivo de este estudio fue analizar hasta qué punto el procesamiento de magnitudes numéricas no simbólicas o simbólicas predice la futura ejecución en matemáticas más allá de factores no numéricos, y si esta predicción depende de la numerosidad de las cantidades. Para ello se utilizaron tareas de comparación (simbólicas y no simbólicas y con cantidades pequeñas y grandes), ya que son las tareas usadas más frecuentemente para evaluar el procesamiento de magnitudes numéricas. Además, de cada tarea se tomaron dos tipos de medidas, unas relacionadas con la eficacia para procesar magnitudes a partir del porcentaje de errores y TR, y otras que reflejan la precisión

---

<sup>(2)</sup> Linsen, Verschaffel, Reynvoet, and De Smedt (2014) también analizaron la relación entre el procesamiento de magnitudes numéricas y ejecución matemática con cantidades pequeñas y grandes, pero su estudio se realizó con participantes de un nivel educativo más alto y se basó en datos transversales.

de, o el acceso a, la representación de la magnitud. Para investigar si la asociación entre procesamiento de magnitudes numéricas y ejecución matemática fue específica, también se administró un test de habilidades cognitivas generales junto con una tarea computarizada de velocidad de procesamiento. Dado que la asociación entre procesamiento de magnitudes y ejecución en matemáticas puede estar mediatizada por el dominio matemático analizado (i.e., por las medidas de ejecución consideradas, que van desde medidas de ejecución matemática general hasta medidas más específicas como por ejemplo el cálculo simple; ver Schneider et al., 2016), en el presente estudio se administraron dos pruebas de ejecución, una estandarizada de aritmética elemental (resolución de problemas y cálculo simple) y una prueba no estandarizada que evalúa flexibilidad en el uso de los números para utilizar el conocimiento conceptual de las propiedades aditivas.

## **Método**

### **Participantes**

En el estudio participaron 55 estudiantes de primer curso de Educación Primaria con una media de edad de 75 meses ( $DT = 4$  meses). Los participantes pertenecieron a dos aulas de un mismo colegio concertado ubicado en una zona de nivel sociocultural medio de la ciudad de Salamanca. Tres participantes fueron excluidos por no completar todas las tareas, por lo que todos los análisis estadísticos presentados más adelante se ejecutaron sobre una muestra de 52 estudiantes (24 niñas y 28 niños). Los padres de los participantes dieron su consentimiento por escrito y se siguieron las normas del Comité de Bioética de la Universidad en la que se desarrolló el estudio.

### **Instrumentos**

#### **Comparación de magnitudes grandes no simbólicas (CMGNS)**

Esta es la tarea más ampliamente utilizada para medir la precisión del SNA. Los participantes tenían que elegir el más numeroso de dos conjuntos

de puntos presentados simultáneamente a ambos lados de una pantalla de ordenador (15 pulgadas) apretando las teclas S (más numeroso a la izquierda) o L (más numeroso a la derecha). El rango de numerosidad fue desde 6 a 60 puntos con tres ratios diferentes (.50, .66 y .75). La prueba constó de 90 ensayos (30 por ratio) divididos en dos bloques más 5 ensayos de práctica. Cada ensayo comenzó con un punto de fijación (1000 ms) seguido por los conjuntos de puntos que permanecieron un tiempo limitado en la pantalla (1500 ms) para evitar conteo. Para prevenir el uso de indicios no numéricos en las respuestas, en la mitad de los ensayos el tamaño de los puntos y el área total correlacionó positivamente con la numerosidad (i.e., a mayor numerosidad mayor tamaño de los puntos y mayor área ocupada) y en la otra mitad correlacionó negativamente (i.e., a mayor numerosidad menos tamaño de los puntos y menor área ocupada).

### **Comparación de magnitudes grandes simbólicas (CMGS)**

En esta tarea los participantes tenían que comparar un número arábigo presentado en el centro de la pantalla de un ordenador con un número de referencia fijo (55), apretando la tecla de la derecha (L) si el número era mayor y la tecla de la izquierda (S) si el número era menor. Se presentaron todos los números entre 31 y 79 (excepto el estándar 55) dos veces, dando un total de 96 estímulos divididos en dos bloques más 5 ensayos de práctica. Cada ensayo fue precedido de un punto de fijación (1000 ms) seguido por el número que permaneció en la pantalla hasta la respuesta del participante.

### **Comparación de magnitudes pequeñas no simbólicas (CMPNS)**

Los participantes tenían que comparar el más numeroso de dos conjuntos de cuadrados presentados simultáneamente a ambos lados de una pantalla de ordenador, apretando las teclas S (más numeroso a la izquierda) o L (más numeroso a la derecha). Se presentaron 72 ensayos (más cinco de práctica) con combinaciones de 1 a 9 cuadrados con distancias entre conjuntos de 1 a 6. Para asegurar que las respuestas no se basaran en indicios no numéricos, el área individual, el área total y

la densidad de los cuadrados se variaron sistemáticamente. Cada ensayo comenzó con un punto de fijación (1000 ms) seguido por los conjuntos de cuadrados que permanecieron en la pantalla hasta la respuesta. Los participantes fueron instruidos a responder tan rápida y exactamente como fuera posible sin contar.

### **Comparación de magnitudes pequeñas simbólicas (CMPS)**

Esta tarea fue similar a la comparación de magnitudes pequeñas no simbólicas sustituyendo los conjuntos de cuadrados por los números arábigos correspondientes a cada numerosidad.

### **Ejecución matemática**

La competencia matemática se evaluó a partir de las pruebas de Cálculo Numérico y Problemas Numérico-Verbales de la Batería de Aptitudes Diferenciales y Generales (BADyG; Yuste, Yuste, Martínez, y Galve, 2012).

Adicionalmente se administró una prueba no estandarizada de resolución de problemas aritméticos a través de cálculo mental (CM en adelante), en la que los participantes necesitaron hacer un uso flexible de los números desde la comprensión de su magnitud para llegar al resultado. La prueba constó de 12 problemas de un paso de las categorías de Cambio 1, 2 y 3 (de acuerdo al esquema de clasificación de Riley, Greeno, y Heller, 1983) con la pregunta al inicio del enunciado. La mitad de los problemas se podían resolver simulando directamente la acción descrita en el problema (Brissiaud y Sander, 2010), como en “Cuántas canicas tiene Pedro ahora, si al principio de una partida tenía 42 canicas y después perdió 3”; en este ejemplo se puede simular la acción restando directamente el sustraendo desde el minuendo. En la otra mitad de los problemas fue necesario hacer uso del conocimiento conceptual de los principios aritméticos, como la conmutatividad o la relación inversa entre suma y resta, para llegar al resultado, como en, “Cuántas canicas tiene Pedro ahora, si al principio de una partida tenía 42 canicas y después perdió 39”; en este caso una simulación directa de la acción implicaría un alto costo cognitivo (Brissiaud y Sander, 2010), mientras que a partir del principio de inversión se puede determinar rápidamente cuánto

hay que añadir al sustraendo para llegar al minuendo. Para asegurar el uso de estas estrategias solo se permitió 6 s para la resolución de cada problema. Por ello, uno de los números (o ambos) fue entre 30 y 70 y el otro (o la distancia entre ambos números grandes) 3 o 4, y la operación siempre implicó cambio de decena.

## **Tareas de control**

### **Habilidad intelectual**

Para controlar la habilidad intelectual se utilizó el Test ICCE de Inteligencia (Yuste, Franco, y Palacios, 2013), una medida de inteligencia general que incluye factores como el verbal, numérico, espacial y razonamiento lógico.

### **Velocidad de procesamiento (VP)**

Esta tarea se incluyó para controlar la velocidad de respuesta con el teclado del ordenador. Se presentaron dos cuadrados (uno negro y el otro rojo) a ambos lados de la pantalla del ordenador, y los participantes tenían que apretar lo más rápido posible la tecla correspondiente al lado de la pantalla donde apareció el cuadrado rojo. La prueba incluyó 20 ensayos.

## **Procedimiento y Diseño**

Todas las tareas computarizadas se aplicaron individualmente a los participantes en una sala aislada de ruidos en el centro escolar. La presentación de los estímulos y la recogida de datos fueron controlados por el software Cedrus SuperLab™ ([www.cedrus.com](http://www.cedrus.com)). La aplicación de estas pruebas se llevó a cabo en el primer curso de Educación Primaria durante los meses de abril y mayo junto con otras tareas no analizadas en este trabajo. Los resultados del test de inteligencia fueron proporcionados por el centro escolar. La ejecución en matemáticas se evaluó en una aplicación colectiva dos años más tarde, en tercero de Educación Primaria. De acuerdo con los objetivos del estudio, y siguiendo la clasificación de Ato, López y Benavente (2013), se realizó un diseño correlacional en el que utilizamos una estrategia asociativa, explorando las relaciones entre las variables estudiadas. Concretamente, se trata de un diseño predictivo

longitudinal, en el que analizamos una única muestra tomando medidas en dos momentos diferentes. Para el análisis de los datos se empleó el programa estadístico SPSS Statistic 23.0. En un primer momento se obtuvieron los datos descriptivos de los 52 participantes en cada una de las tareas de comparación de magnitudes y se verificó la presencia de los efectos de ratio y distancia utilizando análisis de varianza de medidas repetidas. Posteriormente se llevaron a cabo análisis de correlación de Pearson y de regresión jerárquica para analizar los efectos de las variables predictoras sobre las variables criterio.

## Resultados

Para cada tarea de procesamiento de magnitudes se tomaron las siguientes medidas. En la tarea de comparación de magnitudes grandes no simbólicas se tomó como medida de eficacia la proporción de aciertos. Además, se calculó para cada participante la fracción de Weber, un índice de la precisión de la representación de la magnitud no simbólica, o lo que es lo mismo, la agudeza del SNA (ver procedimiento en Halberda y Feigenson, 2008). Una mayor fracción de Weber refleja una menor agudeza del SNA. En la tarea de comparación de magnitudes grandes simbólicas los números se agruparon en 6 distancias (i.e., la distancia 1 incluyó los números 51-54 y 56-59, esto es, 1 a 4 unidades de distancia desde el 55; la distancia 2 incluyó los números 47-50 y 60-63, y así sucesivamente hasta la distancia 6 que incluyó los números 31-34 y 76-79). En cada distancia se calculó un TR ajustado, una medida que combinó velocidad y exactitud, dividiendo el tiempo de respuesta por la proporción de aciertos. A partir de aquí se tomaron dos medidas, una de eficacia (la media de los TR ajustados), y la otra medida reflejo de la precisión de la representación de la magnitud, indicada en el tamaño del efecto de distancia, que se calculó para cada participante a partir de la pendiente de la regresión lineal para los TR ajustados utilizando la distancia como variable predictora. Se asume que la pendiente debería ser negativa, porque a mayor distancia menor TR. Y una mayor pendiente reflejaría una representación de la magnitud menos precisa. Estas mismas medidas de eficacia (media de los TR ajustados) y precisión de la representación de la magnitud (tamaño del efecto de distancia) se calcularon para las dos tareas de comparación de magnitudes pequeñas.

**TABLA I.** Media (y desviación típica) de TR (ms) y aciertos (proporción) de las tareas de comparación de magnitudes

		DISTANCIA					
		1	2	3	4	5	6
CMGS	TR	1731 (443)	1523 (415)	1430 (418)	1419 (425)	1404 (389)	1350 (419)
	AC	.76 (.16)	.82 (.18)	.86 (.16)	.88 (.13)	.88 (.14)	.91 (.10)
CMPS	TR	1284 (323)	1210 (308)	1155 (257)	1164 (310)	1132 (286)	1091 (276)
	AC	.88 (.12)	.90 (.10)	.94 (.06)	.95 (.05)	.95 (.05)	.97 (.03)
CMPNS	TR	1332 (360)	1251 (363)	1148 (287)	1121 (237)	159 (249)	158 (250)
	AC	.80 (.14)	.89 (.11)	.92 (.09)	.95 (.08)	.97 (.07)	.97 (.07)
		RATIO					
		<b>0.75</b>	<b>0.75</b>	<b>0.75</b>			
CMGNS	AC	.76 (.12)	.65 (.10)	.62 (.10)			

*Nota:* CMGS = comparación de magnitudes grandes simbólicas; CMPS = comparación de magnitudes pequeñas simbólicas; CMPNS = comparación de magnitudes pequeñas no simbólicas; CMGNS = comparación de magnitudes grandes no simbólicas.

Para verificar la presencia de los efectos de ratio y distancia en las tareas de comparación de magnitudes se llevaron a cabo análisis de varianza de medidas repetidas, con la ratio (3 niveles: 0.50, 0.66 y 0.75) y distancia (6 niveles: de distancia 1 hasta 6) como factores intra-sujetos. Como se esperaba, en la tarea de comparación de magnitudes grandes no simbólicas hubo un efecto de ratio,  $F(2, 102) = 60.82$ ,  $p < .0001$ ,  $\eta^2 = .54$ , indicando que la proporción de aciertos disminuyó con el aumento de la ratio (ver Tabla 1). De la misma manera, en las demás tareas de comparación de magnitudes hubo un efecto de distancia ( $F(5, 255) = 20.09$ ,  $p < .0001$ ,  $\eta^2 = .28$ ;  $F(5, 255) = 19.31$ ,  $p < .0001$ ,  $\eta^2 = .27$ ; y  $F(5, 255) = 62.06$ ,  $p < .0001$ ,  $\eta^2 = .55$ , para comparación de magnitudes grandes simbólicas, comparación de magnitudes pequeñas simbólicas y comparación de magnitudes pequeñas no simbólicas, respectivamente). Estos efectos de distancia mostraron que los TR ajustados disminuyeron a medida que aumentó la distancia (ver Tabla 1). Estos mismos efectos han sido encontrados sistemáticamente en los estudios previos sobre el tema.

Para analizar la relación entre las distintas medidas de procesamiento de magnitudes y la ejecución en matemáticas se calcularon coeficientes de correlación de Pearson (Tabla 2). No se observó ninguna correlación

significativa entre las medidas de comparación de magnitudes grandes no simbólicas y ejecución en matemáticas. Las demás medidas relacionadas con la comparación de magnitudes correlacionaron significativamente con las dos tareas de ejecución en matemáticas, excepto el índice de eficacia de comparación de magnitudes pequeñas no simbólicas, que no correlacionó con la tarea de CM. Respecto a las demás medidas, una ejecución más alta en matemáticas se asoció con más rápidos TR y un menor tamaño del efecto distancia.

TABLA II. Correlaciones de Pearson entre las distintas variables

	1	2	3	4	5	6	7	8	9	10
1 BADyG.	—	.81**	.08	-.08	-.57**	.45**	-.38**	.38**	-.48**	.34*
2 CM			-.01	-.20	-.57**	.48**	-.20	.35*	-.34*	.35*
3 CMGNS (E)			—	.83**	-.02	-.18	-.24	.19	-.14	-.05
4 CMGNS (w)				—	.09	-.18	-.17	.02	.04	-.17
5 CMGS (E)					—	-.29*	.41**	-.46**	.50**	-.54**
6 CMGS (ID)						—	-.18	.03	-.14	.07
7 CMPNS (E)							—	-.70**	.55**	-.31*
8 CMPNS (ID)								—	-.48**	.48**
9 CMPS (E)									—	-.40**
10 CMPS (ID)										—

Nota: CMGNS = comparación de magnitudes grandes no simbólicas; E = eficacia; w = fracción de Weber; CMGS = comparación de magnitudes grandes simbólicas; ID = índice de distancia; CMPNS = comparación de magnitudes pequeñas no simbólicas; CMPS = comparación de magnitudes pequeñas simbólicas.

\*  $p < .05$  \*  $p < .01$

Por último, para analizar la contribución de cada una de las medidas procesamiento de magnitudes a la varianza en ejecución en matemáticas se llevaron a cabo dos análisis de regresión jerárquica (Tabla 3). Estos análisis fueron calculados para cada medida de ejecución matemática por separado. En el Paso 1 se incluyeron inteligencia y velocidad de procesamiento para controlar las habilidades cognitivas generales. En el Paso 2, se incluyeron las medidas de procesamiento de magnitudes que mostraron una correlación significativa con la ejecución en matemáticas.

La inclusión de todas las medidas en un único bloque permitió distinguir la varianza en ejecución aritmética atribuible a las medidas de procesamiento de magnitudes más allá de la varianza explicada por las habilidades cognitivas generales.

**TABLA III.** Análisis de regresión jerárquica prediciendo las dos medidas de ejecución matemática

BADyG					CM			
Paso	Predictor	$\beta$	t	R <sup>2</sup>	Predictor	$\beta$	t	R <sup>2</sup>
1	Inteligencia	.22	1.81 <sup>a</sup>	.29	Inteligencia	.32	2.69*	.35
	VP	-.07	-.43		VP		.06	.44
2	CMPNS (E)	-.02	-1.10	.52	CMPNS (ID)	.10	.79	.54
	CMPNS (ID)	.09	.55		CMPS (E)	-.10	-.80	
	CMPS (E)	-.26	-1.88 <sup>b</sup>		CMPS (ID)	.06	.48	
	CMPS (ID)	-.01	-.08		CMGS (E)	-.31	-1.77 <sup>c</sup>	
	CMGS (E)	-.20	-1.12		CMGS (ID)	.26	2.22*	
	CMGS (ID)	.26	2.19*					

Nota: VP, velocidad de procesamiento; CMPNS = comparación de magnitudes pequeñas no simbólicas; E = eficacia; ID = índice de distancia; CMPS = comparación de magnitudes pequeñas simbólicas; CMGS = comparación de magnitudes grandes simbólicas;

\*  $p < .05$  <sup>a</sup>  $p = .08$  <sup>b</sup>  $p = .06$  <sup>c</sup>  $p = .08$

En el primer análisis de regresión jerárquica, el modelo completo explicó el 52% de la varianza en el BADyG,  $F(8, 43) = 5.72$ ,  $p < .0001$ , y las medidas de procesamiento de magnitudes añadieron un 23% más allá de las variables de control,  $F(6, 43) = 3.34$ ,  $p < .001$ . No obstante, solo el índice de distancia en la tarea de comparación de magnitudes grandes simbólicas fue un predictor significativo, y la eficacia para comparar magnitudes pequeñas simbólicas se aproximó a la significatividad. En el segundo análisis de regresión jerárquica, el modelo completo explicó el 54% de la varianza de CM,  $F(7, 44) = 7.43$ ,  $p < .0001$ , y las medidas de procesamiento de magnitudes añadieron un 19% a la varianza explicada por las variables de control,  $F(5, 44) = 3.63$ ,  $p < .001$ . En este caso, otra vez el índice de distancia en la comparación de magnitudes grandes simbólicas fue el único predictor significativo, con la eficacia para procesar

magnitudes grandes simbólicas aproximándose a la significatividad. Por último, dado que las medidas correlacionaron entre sí, se llevó a cabo un análisis de colinealidad. Los factores de inflación de la varianza (FIV) se encontraron dentro de un rango aceptable (todos los FIVs  $< 3$ ).

## Discusión y Conclusiones

Investigaciones recientes han mostrado que la habilidad para procesar y representar magnitudes numéricas se relaciona con la ejecución en matemáticas, aunque hay controversia sobre el rol que juega el procesamiento de magnitudes no simbólicas y simbólicas en esta relación. Una limitación de estas investigaciones es que la mayoría de ellas han utilizado tareas que incluyen cantidades pequeñas (de un dígito), especialmente cuando se utilizan tareas de comparación de magnitudes simbólicas. Sin embargo, el conocimiento de las magnitudes numéricas simbólicas se expande evolutivamente desde números pequeños a números mayores (Siegler, 2016), por lo que resulta necesario analizar hasta qué punto existe una relación entre procesamiento de magnitudes numéricas y ejecución en matemáticas cuando se utilizan cantidades mayores, especialmente en el periodo educativo en el que se están aprendiendo estas cantidades. Este estudio se diseñó para analizar esta cuestión. Específicamente, se diseñaron tareas de comparación de magnitudes no simbólicas y simbólicas que incluyeron cantidades pequeñas y grandes para analizar la contribución de cada una de ellas a la futura ejecución en matemáticas. Globalmente, los resultados mostraron que las habilidades para procesar magnitudes simbólicas al inicio de la escolaridad formal se relacionan longitudinalmente con la futura ejecución matemática, y que estas relaciones fueron independientes de habilidades cognitivas generales tales como la habilidad intelectual y la velocidad de procesamiento. No obstante, estas relaciones dependieron de las medidas de procesamiento de magnitudes utilizadas y la ejecución matemática analizada.

Considerando la contribución de las medidas en función del formato no simbólico vs. simbólico, los resultados mostraron que el procesamiento de magnitudes no simbólicas grandes no correlacionó con la ejecución en matemáticas, y aunque las medidas de procesamiento de magnitudes no simbólicas pequeñas correlacionaron significativamente con la

ejecución matemática (excepto la medida de eficacia con CM), estas relaciones se eliminaron una vez controladas las habilidades cognitivas generales y las habilidades de procesamiento numérico simbólico. Estos resultados corroboran los encontrados en otros estudios en los que se ha comprobado que cuando se analiza la asociación entre procesamiento de magnitudes numéricas y ejecución matemática combinando medidas de procesamiento de magnitudes simbólicas y no simbólicas, solo las simbólicas predicen la varianza en ejecución matemática (e.g., Lyons, Price, Vaessen, Blomert, y Ansari, 2014; Sasanguie, Göbel, Moll, Smets, y Reynvoet, 2013). Sin embargo, los resultados del presente estudio van más allá de los encontrados en los estudios previos al demostrar que el predictor más robusto de la futura ejecución en matemáticas fue la comparación de magnitudes simbólicas grandes. Esto es consistente con el trabajo de Xenidou-Dervou et al. (2016), quienes encontraron resultados similares en su estudio longitudinal. Pero aun así el presente estudio también añade nueva información, ya que no fue tanto la medida de eficacia para procesar magnitudes simbólicas grandes, sino el tamaño del efecto de distancia, lo que mejor predijo la futura ejecución matemática, indicando que un menor tamaño del efecto de distancia se asoció con una ejecución matemática más alta. Dado que se asume que un menor efecto distancia refleja proyecciones más exactas entre los números arábigos y las magnitudes que representan, estos resultados sugieren que contar con buenas conexiones entre los números y sus significados juega un rol importante en el aprendizaje de las matemáticas. Es interesante notar que Linsen, Verschaffel, Reynvoet, y De Smedt (2014) también encontraron una asociación entre el procesamiento de magnitudes numéricas y la ejecución matemática utilizando tareas que incluyeron magnitudes grandes, pero solo en las medidas de eficacia (i.e., aciertos y TR), mientras que esta asociación no fue evidente cuando consideraron el tamaño del efecto de distancia. No obstante, este estudio se basó en datos transversales, y lo que es más importante, se realizó con participantes de un nivel educativo más alto (tercer curso) que los participantes del presente estudio (primer curso). Esto sugeriría que cuando hay un mayor conocimiento de los números de dos cifras, esto es, representaciones más exactas de los números simbólicos, la eficacia procesando esos números se asociaría con la ejecución matemática. Los niños más jóvenes de primer curso, sin embargo, están comenzando a proyectar los números de dos dígitos en las representaciones de la magnitud, lo que podría explicar

por qué el acceso a esas representaciones es más sensible a predecir la ejecución matemática.

Por lo que se refiere a los resultados obtenidos en función de la ejecución matemática evaluada, aunque estos fueron muy similares, posiblemente debido a la alta correlación positiva entre ambas medidas de ejecución (ver Tabla 2), también se encontró alguna diferencia a considerar. A pesar de que el mejor predictor en ambas medidas de ejecución fue el tamaño del efecto de distancia en la comparación de magnitudes simbólicas grandes, la eficacia para comparar magnitudes simbólicas pequeñas mostró una tendencia a la significatividad ( $p = .06$ ), pero solo en el BADyG. Una posible explicación se puede encontrar en el hecho de que este test evalúa aritmética con tareas que incluyen fundamentalmente números de un dígito, especialmente en el cálculo numérico. Esto es consistente con un reciente estudio en el que se demostró, con participantes que estaban comenzando la escolaridad formal, que la eficacia para comparar magnitudes simbólicas pequeñas (con una tarea idéntica a la del presente estudio) predijo la futura ejecución en una tarea de cálculo con operaciones de un dígito (Vanbinst, Ghesquière, y De Smedt, 2015). La tarea de CM, sin embargo, es una medida de ejecución en la que cobra más relevancia el procesamiento de números de dos dígitos, de ahí que las únicas medidas que predijeron fueron las de comparación de magnitudes simbólicas grandes, posiblemente debido a que un uso flexible de los principios aritméticos, como el principio de inversión, requiere una buena comprensión de las magnitudes de los números involucrados (Linsen et al., 2014).

El presente estudio tiene limitaciones motivadas principalmente por el tamaño de la muestra, lo que hace que las conclusiones deban ser tomadas con cautela. No obstante, la muestra fue bastante homogénea, al ser seleccionada de dos clases un mismo centro con métodos de enseñanza y currículos idénticos, y los resultados consistentes con los encontrados en estudios previos. Aún con esta limitación en mente, los resultados obtenidos tienen algunas implicaciones educativas. Dado que una buena comprensión de los números en términos de sus magnitudes puede ser útil para la futura ejecución en matemáticas, sería interesante desarrollar instrumentos de evaluación que permitieran analizar las diferencias individuales en esta competencia básica, así como identificar tempranamente estudiantes en riesgo de presentar dificultades en el aprendizaje de las matemáticas, algo prioritario en cualquier sistema

educativo desde el punto de vista de la prevención. En cierto sentido, este planteamiento se inspira en el tremendo desarrollo llevado a cabo en el dominio de la lectura, donde otra competencia básica como el procesamiento fonológico subyace a las diferencias individuales en el aprendizaje de la lectura (Melby-Lervag, Lyster, y Hulme, 2012), y juega un importante rol en la detección de estudiantes en riesgo, lo que ha llevado a diseñar test que incluyen medidas relacionadas con esta importante competencia. En el caso de las matemáticas, y a pesar de la importancia demostrada por el procesamiento de la magnitud numérica en relación a la ejecución matemática, los actuales test no incluyen esta competencia entre los aspectos evaluados, y cuando lo hacen, el número de ítems es relativamente escaso y sin condiciones cronometradas. Esto último es importante, porque lo que se relaciona con la ejecución matemática no es tanto procesar magnitudes en sí mismo, sino la rapidez para acceder a esas magnitudes bajo condiciones de tiempo limitado. No obstante, las medidas desarrolladas en los estudios previos (incluyendo el presente estudio) para analizar las habilidades de procesamiento de magnitudes se han basado en tareas aplicadas con ordenador, las cuales permiten recoger el tiempo de respuesta pero que no siempre son adecuadas para su uso en contextos educativos, especialmente por el tiempo de aplicación que consumen y la necesidad de contar con software especializado. En este contexto, Nosworthy, Bugden, Archibald, Evans, y Ansari (2013) diseñaron una tarea rápida (2 minutos) de lápiz y papel para evaluar la habilidad para comparar magnitudes con cantidades pequeñas (de 1 a 9) bajo restricción de tiempo (i.e., número de ítems respondidos correctamente en un periodo de tiempo), y comprobaron que la ejecución en esta tarea se relacionó con la ejecución en matemáticas. Los resultados del presente estudio indicarían que se podrían diseñar tareas similares incluyendo cantidades grandes de dos dígitos. Estas tareas tienen la ventaja de la rápida aplicación dentro del aula y con una mínima instrucción, lo que permitiría la detección temprana de estudiantes en riesgo (Brankaer, Ghesquière, y De Smedt, 2017).

Además de las implicaciones relacionadas con la evaluación, los resultados del presente estudio también tienen implicaciones para la práctica educativa, ya que, si una adecuada comprensión del significado de los números predice la ejecución matemática, sería importante orientar una parte de la enseñanza a reforzar las habilidades de procesamiento de magnitudes numéricas de los aprendices como preparación para

el aprendizaje de contenidos matemáticos más complejos. Aunque la investigación en este campo aún no ha aportado resultados concluyentes (De Smedt et al., 2013; Schneider et al., en prensa), algunos currículos de matemáticas han comenzado a incorporar actividades relacionando números con magnitudes como parte de la enseñanza inicial de las matemáticas. Como ejemplo paradigmático de esta propuesta es el proyecto editorial *My Pals are here*<sup>3</sup> de Singapur, en el que se refuerza la habilidad de los estudiantes para procesar magnitudes numéricas a través de actividades que conectan símbolos numéricos con sus significados.

## Referencias

- Ashkenazi, S., Mark-Zigdon, N., y Henik, A. (2009). Numerical distance, effect in developmental dyscalculia. *Cognitive Development*, 24, 387-400.
- Ato, M., & López, J., & Benavente, A. (2013). Un sistema de clasificación de los diseños de investigación en psicología. *Anales de Psicología*, 29, 1038-1059.
- Barth, H., Starr, A., y Sullivan, J. (2009). Children's mappings of large number words to numerosities. *Cognitive Development*, 24, 248-264.
- Bertelletti, I., Lucangeli, D., Piazza, M., Dehaene, S., y Zorzi, M. (2010). Numerical estimation in preschoolers. *Developmental Psychology*, 46, 545-551.
- Brankaer, C., Ghesquière, P., y De Smedt, B. (2017). Symbolic magnitude processing in elementary school children: A group administered paper-and-pencil measure (SYMP Test). *Behavior Research Methods*, 49, 1361-1373.
- Brissiaud, R., y Sander, E. (2010). Arithmetic word problem solving: a Situation Strategy First framework. *Developmental Science*, 13, 92-107.
- Chen, Q., y Li, J. (2014). Association between individual differences in nonsymbolic number acuity and math performance: A meta-analysis. *Acta Psychologica*, 148, 163-172.

---

<sup>3</sup> Este Proyecto ha sido desarrollado por la editorial Marshall Cavendish para los libros de texto más utilizados en Singapur.

- Dehaene, S. (2011). *The number sense: How the mind creates mathematics*. New York: Oxford University Press.
- De Smedt, B., Noël, M. P., Gilmore, C., y Ansari, D. (2013). How do symbolic and nonsymbolic numerical magnitude processing skills relate to individual differences in children's mathematical skills? A review of evidence from brain and behavior. *Trends in Neuroscience and Education*, 2, 48–55.
- Fazio, L. K., Bailey, D. H., Thompson, C. A., y Siegler, R. S. (2014). Relations of different types of numerical magnitude representations to each other and to mathematics achievement. *Journal of Experimental Child Psychology*, 123, 53–72.
- Feigenson, L., Dehaene, S., y Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences*, 8, 307–314.
- Halberda, J., y Feigenson, L. (2008). Developmental change in the acuity of the 'number sense': The approximate number system in 3-, 4-, 5-, and 6-year-olds and adults. *Developmental Psychology*, 44, 1457–1465.
- Lefevre, J-O., Wells, E., y Sowinski, C. (2016). Individual differences in basic arithmetical processes in children and adults. En R. Cohen Kadosh y A. Dowker (Eds.), *The Oxford handbook of numerical cognition* (pp. 895-914). Oxford: Oxford University Press.
- Leibovich, T., y Ansari, D. (2016). The symbol-grounding problem in numerical cognition: A review of theory, evidence, and outstanding questions. *Canadian Journal of Experimental Psychology*, 70, 12-23.
- Libertus, M. E., Feigenson, L., y Halberda, J. (2011). Preschool acuity of the approximate number system correlates with school math ability. *Developmental Science*, 14, 1292-1300.
- Lipton, J. S., y Spelke, E. S. (2005). Preschool children's mapping of number words to nonsymbolic numerosities. *Child Development*, 76, 978-988.
- Linsen, S., Verschaffel, L., Reynvoet, B., y De Smedt, B. (2014). The association between Children's numerical magnitude processing and mental multi-digit subtraction. *Acta Psychologica*, 145, 75-83.
- Lyons, I. M., y Ansari, D. (2015). Foundations of Children's Numerical and Mathematical Skills: The Roles of Symbolic and Nonsymbolic Representations of Numerical Magnitude. En J. B. Benson (Ed.), *Advances in Child Development and Behavior*, Vol. 48, (pp. 93-116). Burlington: Academic Press.

- Lyons, I. M., Price, G. R., Vaessen, A., Blomert, L., y Ansari, D. (2014). Numerical predictors of arithmetic success in grades 1-6. *Developmental Science*, 17, 714-726.
- Melby-Lervåg, M., Lyster, S. H., y Hulme, C. (2012) Phonological skills and their role in learning to read: A Meta-Analytic Review. *Psychological Bulletin*, 138, 322-352.
- Mundy, E., y Gilmore, C. K. (2009). Children's mapping between symbolic and nonsymbolic representations of number. *Journal of Experimental Child Psychology*, 103, 490-502.
- Nosworthy, N., Bugden, S., Archibald, Evans, B. y Ansari, D. (2013). A two-minute paper-and-pencil test of symbolic and nonsymbolic numerical magnitude processing explains variability in primary school children's arithmetic competence. *PLoS ONE*, 8(7), e67918.
- Orrantia, J., San Romualdo, S., Matilla, L., Sánchez, R., Muñoz, D. y Verschaffel, L. (2017). Marcadores nucleares de la competencia aritmética en preescolares. *Psychology, Society, & Education*, 9, 121-124.
- Reynvoet, B., y Sasanguie, D. (2016). The symbol grounding problem revisited: a thorough evaluation of the ANS mapping account and the proposal of an alternative account based on symbol-symbol associations. *Frontiers in Psychology*, 7:1581.
- Riley, N. S., Greeno, J., y Heller, J. I. (1983). Development of children's problem solving ability in arithmetic. In H. P. Ginsburg (Ed.), *The development of mathematical thinking* (pp. 153-196). New York: Academic Press.
- Sasanguie, D., Göbel, S. M., Moll, K., Smets, K., y Reynvoet, B. (2013). Approximate number sense, symbolic number processing, or number-space mappings: What underlies mathematics achievement? *Journal of Experimental Child Psychology*, 114, 418-431.
- Schneider, M., Beeres, K., Coban, L., Merz, S., Susan Schmidt, S., Stricker, J., y De Smedt, B. (en prensa). Associations of non-symbolic and symbolic numerical magnitude processing with mathematical competence: a meta-analysis. *Developmental Science*. doi:10.1111/desc.12372.
- Siegler, R. S. (2016). Magnitude knowledge: the common core of numerical development. *Developmental Science*, 19, 341-361.
- Vanbinst, K., Ghesquière, P., y De Smedt, B. (2015). Does numerical processing uniquely predict first graders' future development of single-digit arithmetic? *Learning and Individual Differences*, 37, 153-160.

- Xenidou-Dervou, I., Molenaar, D., Ansari, D., van der Schoot, M., y van Lieshout, E. C. D. M. (en prensa). Nonsymbolic and symbolic magnitude comparison skills as longitudinal predictors of mathematical achievement. *Learning and Instruction*. <http://dx.doi.org/10.1016/j.learninstruc.2016.11.001>
- Yuste, C., Franco, J., y Palacios, J. M. (2013). *Test ICCE de Inteligencia*. Madrid: ICCE Publicaciones.
- Yuste, C., Yuste, D., Martínez, R. y Galve, J. L. (2012). *BADyG. Batería de Aptitudes Generales y Diferenciales*. Madrid: CEPE.

**Información de contacto:** Josetxu Orrantia. Universidad de Salamanca. Facultad de Educación. Departamento de Psicología Evolutiva y de la Educación. Pº de Canalejas, 169. 37008 Salamanca. E-mail: orrantia @usal.es

# Analfabetismo en el primer tercio del siglo XX: la creación de escuelas en Valdés (Asturias)

## Illiteracy in the first third of the twentieth century: the creation of schools in Valdés (Asturias)

DOI: 10.4438/1988-592X-RE-2017-381-384

Patricia Suárez

Trinidad Suárez

*Universidad de Oviedo*

### **Resumen:**

En este artículo se analiza la creación de escuelas en el municipio de Valdés (Asturias) durante el primer tercio del siglo XX. Cabe señalar el papel protagonista que tuvieron los vecinos en la materialización del acceso a la enseñanza de sus hijos. Fueron los propios vecinos quienes sufragaron parte de los gastos de construcción de los edificios, e incluso donaron los solares donde se habrían de levantar los mismos, así como los materiales utilizados. El periodo que abarca la Segunda República fue el más fructífero en cuanto a levantamiento de edificios escolares, de la mano de su alcalde Luis Ochoa de Albornoz, que puso gran empeño en atender las demandas de los vecinos canalizadas a través de los Celadores y Alcaldes de Barrio.

*Palabras clave:* analfabetismo, escuelas, Asturias, Segunda República, ciudadanos.

### **Abstract:**

This article analyzes the creation of schools in the municipality of Valdés (Asturias) during the first third of the 20th century. It is worth noting the protagonist role played by residents in the materialization of access to education for their children. The residents themselves incurred part of the construction

costs of the buildings and even donated the lots where they would be raised and the materials used. The period covered by the Second Republic was the most fruitful in terms of erecting school buildings, at the hands of its mayor Luis Ochoa de Albornoz, who put great effort into meeting the demands of the residents channeled through the Neighborhood Leaders and Watchmen.

*Keywords:* illiteracy, schools, Asturias, Second Republic, citizens.

## Introducción

El objetivo fundamental del trabajo es vincular el proceso de institucionalización de la educación con el advenimiento de la Segunda República en España. El devenir histórico es imprescindible para su comprensión puesto que en este contexto el tema educativo pasó a ocupar un lugar preponderante que se materializó en la creación de escuelas en entornos geográficos remotos e inaccesibles. Si algo se puede afirmar con rotundidad es que la población rural española fue la gran olvidada, pese a ser la más numerosa hasta bien entrado el siglo XX. Y quizás el vacío educativo fue el más gravoso de todos cuantos tuvo que soportar, y causa de la mayoría de sus males puesto que las generaciones se sucedían sin vislumbrar apenas atisbos de cultura. Según Bartolomé Cossío, quizás el pedagogo español más eminente a fines del siglo XIX y principios del XX, lo fundamental es “enseñar al niño a ver desde el primer día” (Bartolomé Cossío, 1985, p .49), de ahí la importancia de los maestros y las escuelas. Cuanto más débil es la formación de los progenitores, mayor es el sentido que tiene el desarrollo de una red de escuelas que llene ese vacío y estimule el desarrollo de esos niños y de las generaciones venideras.

Las escuelas como tales apenas existían, ocupando físicamente los peores lugares. Utilizando palabras casi literales de Fernández Soria y Agulló Díaz (2004), las escuelas se ubicaban en ruinosas viviendas, en sótanos o en antiguos establos. Por si fuese poco, la escuela rural era unitaria la mayor parte de las veces, pues la matrícula y los medios eran insuficientes para conformar una escuela graduada, reservada al ámbito urbano. Este hecho suponía un hándicap para el docente, que se ocupaba de un número elevado de alumnos de diversas edades en un

mismo espacio, con todo lo que ello comportaba. Como cabría esperar, el mobiliario y el material didáctico iba parejo al panorama de miseria imperante.

De todo lo anterior se desprende la alta movilidad de los maestros destinados al ámbito rural, pues en cuanto podían ponían rumbo a las ciudades o por lo menos a poblaciones de más tamaño, en donde por lo menos existiese una escuela graduada que cumplierse con sus expectativas y les ofreciese mayores comodidades.

Cabe preguntarse en qué momento la familia es desplazada de forma definitiva por la escuela a la hora de dotar a los niños de una formación generalista que garantizase su alfabetización, es decir, el control de las técnicas elementales de la lectura, la escritura y el conteo. Los economistas consideran que la educación es una inversión que permite al individuo mejorar su capacidad futura para producir y generar una mayor ganancia y, por lo tanto, afectará positivamente a su bienestar. En este sentido, los padres manifiestan anhelar la prosperidad y grandeza del país a través de una mayor cultura de sus hijos/as en las múltiples solicitudes que cursaron para pedir la puesta en marcha de escuelas a principios del siglo XX (véase apéndice).

El núcleo empírico de este trabajo se centra en Valdés, un municipio del Principado de Asturias. En esta región, a finales del siglo XIX e inicios del XX, el dominio de las tres reglas (leer, escribir y contar) se adquiría fundamentalmente en el entorno familiar (entre quienes se lo podían permitir) o en alguna de las pocas escuelas que había. En este contexto el aporte de los emigrantes a América fue decisivo para impulsar la construcción de escuelas y erradicar el analfabetismo<sup>1</sup>.

La estructura del artículo es la siguiente, en primer lugar se analizan los antecedentes históricos de las leyes educativas más importantes, que nos permiten entender la situación en cuanto a escasez de escuelas en la primera parte del siglo XX.

En el segundo apartado se analizan los fondos y archivos disponibles para el estudio de las escuelas que se construyeron en el primer tercio del siglo XX: información procedente del Archivo Histórico de Asturias y del Archivo municipal de Valdés. El estudio empírico realizado se centra en los orígenes de las escuelas de Valdés, la mayor parte de ellas levantadas

---

<sup>1</sup> Un análisis detallado sobre esta cuestión se puede encontrar en el Informe realizado por el Inspector de Primera Enseñanza de Oviedo, D. Benito Castrillo Sagredo (1926), disponible en la Biblioteca Virtual del Principado de Asturias.

siendo Alcalde Luis Ochoa de Albornoz que dio curso a las demandas de los vecinos del concejo, que deseaban la alfabetización de sus hijos. Por último, el trabajo finaliza con un apartado de conclusiones.

## I. Algunos antecedentes históricos

Por ser un aspecto importante en la configuración de la realidad que nos ocupa, es pertinente un repaso de la legislación gestada a lo largo del periodo histórico al que tratamos de aproximarnos, teniendo en cuenta que su mera publicación no significó en muchas ocasiones su cumplimiento, lo que originó sin duda un desfase entre lo proyectado y la realidad.

Los primeros pasos en la regulación de la escuela pública en España se dieron en época de Carlos III con la promulgación de la Real Provisión de 11 de julio de 1771, iniciándose una tímida legislación en cuanto a la conveniencia de la educación para los niños y niñas, las exigencias para ejercer como maestro, libros escolares, etcétera.

El siguiente hito ya sería la Constitución de 1812, en cuyo título IX, artículo 366, se recoge:

*“En todos los pueblos de la monarquía se establecerán escuelas de primeras letras, en las que se enseñará a los niños a leer, escribir y contar y el catecismo de la religión católica...”*

Tortella y Núñez (2011) hacen hincapié en la importancia del *Informe Quintana*, que reflejó los preceptos constitucionales de 1812 y las ideas de reconocidos ilustrados como Jovellanos. Manuel José Quintana presidió la comisión que emitió el mencionado Informe en 1813, designando al Estado como responsable de la enseñanza y estableciendo el principio de generalidad y gratuidad. A pesar de ello, habría que esperar hasta el año 1857 para que viese la luz la primera Ley General de Instrucción Pública en España, conocida como Ley Moyano, por ser este el ministro encargado de propulsarla y que tomó como referencia el mencionado Informe Quintana.

La Ley Moyano estableció los tres niveles que aún hoy en día, con ciertas modificaciones, perduran: elemental, bachillerato y superior o universitaria. Establece la obligatoriedad de una primera enseñanza

elemental que iría de los seis a los nueve años, cursada en escuelas segregadas por sexos que se ubicarían en las poblaciones de más de 500 almas. La ley señala que la gratuidad sólo sería para aquellos que no tuviesen medios para pagar, lo que en la práctica se traducía en una marginación y etiquetaje de los niños pobres. Además, el principio de generalidad queda limitado a las escasas escuelas existentes. El funcionamiento de las escuelas de enseñanza elemental quedaba en manos de las Juntas locales de primera enseñanza y en virtud del Concordato de 1851, la Iglesia católica quedaba legitimada para intervenir activamente en la educación, especialmente se admite la corresponsabilidad en la enseñanza secundaria.

La configuración de instituciones educativas de carácter estatal comenzó por los estudios superiores, que servirán de punta de lanza para el desarrollo a lo largo del siglo XIX de la escuela primaria, vinculada a la acción del Estado como medio para reforzar su unificación (Egido Gálvez, 1994). Este hecho también es subrayado por Núñez (2005) quien además destaca la importancia que tuvo la creación del Ministerio de Instrucción Pública y Bellas Artes en 1900 como cénit de la implantación de la instrucción primaria.

Llegados a la década de los veinte del pasado siglo la situación era alarmante; sesenta años después de la promulgación de la Ley Moyano la realidad es que apenas hay escuelas para la mitad de la población en edad escolar. López Martín (1997) sentencia que el incremento en la creación de escuelas en el primer tercio del siglo XX no eliminó los problemas de analfabetismo en España y que la red de escuelas seguía siendo exigua e incidental. A modo de ejemplo, en 1900 el porcentaje de analfabetos en España es de aproximadamente el 65% de la población y en 1920 del 52%<sup>2</sup>.

Así pues, el régimen primorriverista se marca como uno de sus objetivos en materia educativa la construcción de escuelas, aunque sin olvidar su papel como vehículo de expansión de los valores ideológicos de la dictadura. Aun existiendo controversia en cuanto al número total de escuelas levantadas en el período que va de 1923 a 1930, parece que no fueron menos de 6.000, siendo loable el esfuerzo aunque la situación deficitaria siguiera manteniéndose.

---

<sup>(2)</sup> Véase López Martín (1994).

López Martín (1994) analiza detalladamente la educación en la dictadura de Primo de Rivera y señala que este empeño constructivo se vio favorecido por la relativa bonanza económica de los años veinte fruto en su mayor parte de la neutralidad de España durante la Gran Guerra. A modo de ejemplo, según el Anuario Estadístico de España citado en López Martín (1994), en la provincia de Oviedo se estimaba en 1920 un total de 118.920 habitantes en edad escolar, para un total de 1.327 escuelas en el año 1923, con una población escolar similar a provincias como Madrid, Barcelona o Valencia, conformando el cuarteto que sobrepasa la barrera de los 100.000.

En 1923 Asturias era la provincia española que contaba con un mayor número de escuelas en términos absolutos, alcanzando una ratio de casi 90 alumnos por centro. Este hecho se debe fundamentalmente a su particularísima configuración geográfica y poblacional y a la valiosa aportación económica de los emigrantes a América. Asturias fue pionera en la industrialización del país gracias a las cuencas carboníferas de Mieres y Langreo, lo que generó un aumento de la demanda educativa sin precedentes.

Con el advenimiento de la Segunda República la educación pasó a ocupar un lugar preponderante, promovándose la creación de numerosas escuelas, pues se estimaba que más de un millón de niños y niñas estaban sin escolarizar. Son elocuentes las palabras del Catedrático de Historia Contemporánea de la Universidad de Salamanca Francisco de Luis, *“la educación es la base de la democracia que quiere instaurar la República”*<sup>3</sup>. Se puso en marcha un Plan quinquenal de construcciones escolares nunca antes visto, llegando a edificarse unas 16.000 de las 27.000 proyectadas.

## 2. Las escuelas en asturias

Hernández Díaz (2000) en su trabajo sobre la escuela rural señala la escasez de recursos de los ayuntamientos, especialmente en la España de posguerra, y la necesaria prestación directa de los campesinos en la construcción de nuevas escuelas. Además, con la llegada de Franco al poder,

---

<sup>3</sup> Véase documental titulado “La República de los maestros” <https://www.youtube.com/watch?v=JVjommxbxQSU>

una vez consolidado el régimen, se publica en 1945 la Ley de Educación Primaria, en la que la Iglesia juega un papel preponderante, recogién-dose en la misma “el principio cristiano de la separación de sexos en la enseñanza”. La Ley establecía cuatro períodos de graduación escolar: 1) período de iniciación, que comprendía las escuelas maternas hasta los cuatro años y las escuelas de párvulos, de los cuatro a los seis, 2) período de enseñanza elemental, de los seis a los diez años, y 3) período de perfeccionamiento, de los diez a los doce años. Éstas dos etapas eran las estrictamente obligatorias (seis a doce años). Por último, el período de iniciación profesional, de los doce a los quince años. Puelles Benítez (1991) también remarca el carácter confesional de la enseñanza así como la dejadez de la dictadura a la hora de promover la creación de nuevas escuelas a pesar del déficit existente.

La radiografía escolar asturiana de la primera mitad del siglo XIX se asentaba sobre las escuelas de primeras letras heredadas del Antiguo régimen, arraigadas en la comunidad vecinal y dependientes económicamente de recursos variados, algunas con propiedades vinculadas, otras sostenidas por las Juntas vecinales, dependientes de Fundaciones o de la Iglesia (Mato Díaz, 2010).

En la Asturias rural, y no olvidemos que en esa época la mayoría de la población se asentaba en este ámbito de una forma muy dispersa, lo normal era que las escuelas estuvieran sostenidas por las Juntas vecinales y atendidas por maestros carentes de titulación, y en el caso de aldeas pequeñas y aisladas la instrucción se centraba en un período pequeño del año, que coincidía con el tiempo en que no se realizaban labores importantes en el campo, dando lugar a las escuelas temporeras, que representaban una cuarta parte del total según Mato Díaz (1992).

Sería la Ley Moyano de 1857 la que concretó el control de la enseñanza por parte del Estado, estableciendo la dotación de una escuela de niños y otra de niñas cada quinientos habitantes, asimismo designaba a los ayuntamientos como los responsables del mantenimiento de las escuelas. Pero en Asturias una aplicación rigurosa de la ley obligaría a la implantación del doble de las escuelas que ya existían, por lo que a través del Arreglo especial de 21 de febrero de 1860 se dispensó a los municipios de tal responsabilidad, perpetuándose así la situación heredada, que significaba la práctica desescolarización de un amplio sector de la población infantil.

Respecto a la formación de los maestros, aunque ya desde 1844 existe escuela Normal de Maestros en Oviedo, los alumnos egresados eran escasos y no cubrían la totalidad de plazas existentes, apenas sólo las de las Escuelas Primarias Superiores y las de las Completas, quedando el resto en manos de maestros habilitados en el mejor de los casos y de maestros sin título, los conocidos como “maestros de medio año”. Luis Bello relata en su obra *Viaje por las escuelas de España* (1926), que estos “maestros” contratados por los vecinos de los pueblos que no contaban con escuela, comían cada día en una casa, de ahí la denominación de “catapotes”, y “babianos” porque en su mayoría provenían de Omaña y Babia, comarcas del noroeste de León lindantes con Asturias, pues eran apreciados porque hablaban el castellano correctamente mientras que aquí los niños de zonas muy rurales solían tener como lengua materna los distintos bables<sup>4</sup>. Precisamente, Soler Mata (2008, p. 15) señala la necesidad de que maestros con titulación ocupen las escuelas del ámbito rural de tal forma que la imagen de la escuela sea el “rostro del maestro”.

En Gera (Tineo-Asturias) se celebra por Todos los Santos una renombrada feria de ganado en la que los vecinos de los pueblos aprovechaban para contratar maestros, de ahí el uso del término “maestros de medio año”, puesto que se ponía escuela desde noviembre al primero de mayo, dejando la época estival para las labores del campo, en las que la mano de obra infantil era fundamental.

Hecha la revisión legislativa acerca del tema que nos ocupa, el resto del apartado pretende reflejar la situación de la educación en un municipio en particular, Valdés. El apartado se estructura de la siguiente forma: se explica la procedencia de la información recabada así como la utilidad de los archivos para la historiografía educativa. A continuación, se explota de forma detallada la documentación reunida con el objetivo de esclarecer el origen de las escuelas valdesanas. Finalmente se hace referencia a las escuelas parvulistas del municipio de Valdés.

## 2.1. Fondo de inspección educativa y estadística de la enseñanza primaria

La principal fuente utilizada en este artículo han sido los fondos del Archivo Provincial de Asturias en los que existe información relativa a

---

<sup>(4)</sup> Véase Bello (1926).

la enseñanza recabada por el Servicio de Inspección de la Consejería de Educación<sup>5</sup>. El Instituto Nacional de Estadística elaboraba una estadística a nivel nacional denominada “Estadística de la Enseñanza Primaria” que recogía información sobre el centro, el profesorado, los alumnos, así como la enseñanza de adultos.

En cuanto a los datos del centro se recogía información sobre el régimen del mismo, si era estatal o no, dentro de los estatales los había de régimen ordinario de provisión o de régimen especial (Patronato), y los no estatales podían depender bien de la Iglesia o ser privados.

Respecto a la clase de centro se distingue entre:

- Escuela unitaria, que era la escuela de un solo maestro, serían aquellas en las que un único maestro tenía a su cargo todas las enseñanzas de los distintos cursos, de acuerdo con el período de escolaridad obligatoria. Podían ser de niños, de niñas o mixtas, éstas últimas se dieron en aquellas zonas en las que no había matrícula suficiente para dotar dos escuelas.
- Escuela graduada sería la entidad docente formada por un número variable de unidades (nunca más de ocho) cuyos maestros tenían a su cargo las enseñanzas de los distintos cursos de escolaridad primaria obligatoria.
- Grupo escolar era la entidad docente primaria constituida por ocho o más unidades escolares, para alumnos de un mismo sexo, con “Director sin curso”.

Respecto a los datos del profesorado se recogía información sobre si había directores sin grado y número de maestros, divididos por sexos.

En relación a los datos de los alumnos, se cubría una plantilla que mostraba el número de alumnos/as según la edad, la asistencia media desglosada por sexos, el número de certificados expedidos en el curso anterior, bien de estudios primarios o de escolaridad, el número de alumnos internos, mediopensionistas (que hacen al menos una comida facilitada por el propio centro) y gratuitos (que no pagan cuota o pensión por la jornada normal de clase).

Por último se recoge información sobre la enseñanza de adultos, es decir, sobre el número de alumnos y alumnas matriculados, la asistencia

---

<sup>5</sup> Información procedente de las fichas de identificación de escuelas 1958-1965. Caja 11292.

media y los que no sabían leer ni escribir al comenzar el curso. Por ejemplo en Oviedo, según informe elaborado por el Inspector Jefe en 1970, se contabilizan 3.015 niños y niñas sin escolarizar o muy deficientemente escolarizados, sobre un total de 18.453 de población en edad de acudir a la escuela. Estas cifras llaman poderosamente la atención por tratarse precisamente de un ayuntamiento como el de Oviedo, capital de la provincia, y bien comunicado si lo comparamos con otras zonas de la región, y por darse en una fecha tan reciente, lo que nos lleva a pensar en la situación que sin duda se vivió en fechas más alejadas y en parajes menos urbanizados y más remotos e inaccesibles de Asturias.

## 2.2. Las escuelas en el concejo de Valdés

A finales de los sesenta, según la Estadística de Enseñanza Primaria mencionada en la sección anterior, el número de escuelas en el concejo de Valdés ascendía a 95 siendo 31 de ellas escuelas mixtas, 28 unitarias de niñas, 28 unitarias de niños, seis de párvulos, y una graduada de niñas y otra de niños ubicadas estas últimas en Luarca, capital del concejo.

Mientras que esta sección se dedica única y exclusivamente a analizar la distribución de las Escuelas que existían a finales de los años sesenta en el concejo de Valdés por pueblo y parroquia, en la sección siguiente se analizan los orígenes históricos de algunas de esas escuelas a través de la documentación encontrada en los fondos del Ayuntamiento de Valdés.

La estructura ocupacional de la población del municipio en 1960 mostraba las características propias de una sociedad todavía agrícola, con un 70% de su población activa dedicada a la actividad agraria. El peso de esta actividad dentro de los municipios rurales es del 54%, de manera que a pesar de ser Luarca un municipio de más de 10.000 habitantes, clasificado como zona urbana intermedia, su estructura económica era fundamentalmente rural. A modo de resumen, en la Tabla 1 se recoge la distribución de la población activa por grupos de actividad en la provincia, Oviedo y Luarca.

TABLA I. Población activa por grupos de actividad económica.

	Provincia	Oviedo	Luarca	Provincia %	Oviedo %	Luarca %
Población activa TOTAL	391.117	40.621	9.572	-	-	-
Agricultura, silvicultura, caza y pesca	151.665	1.739	6.767	38,8	4,3	70,7
Explotación de minas y canteras	54.854	1.388	25	14,0	3,4	0,3
Industrias fabriles	65.463	8.947	764	16,7	22,0	8,0
Construcción	23.428	4.726	615	6,0	11,6	6,4
Electricidad, agua, gas y servicios Sanitarios	2.670	609	33	0,7	1,5	0,3
Comercio	22.905	5.242	336	5,9	12,9	3,5
Transportes, almacenaje y comunicaciones	17.136	2.991	177	4,4	7,4	1,8
Servicios oficiales públicos y con personas	40.575	13.508	533	10,4	33,3	5,6
Actividades mal especificadas	12.421	1.471	322	3,2	3,6	3,4

Fuente: Elaboración propia a partir de los datos del Censo de 1960 (Instituto Nacional de Estadística).

A la vista de los datos ofrecidos en la Tabla 1 parece razonable suponer que la estructura ocupacional en la primera mitad del siglo XX era eminentemente rural. La baja tasa de urbanización unida a unas muy deficientes infraestructuras justificaba la dotación de escuelas en los principales pueblos del concejo.

La mayor parte de los niños y niñas acudían a la escuela andando varios kilómetros, de manera que la ausencia de ésta en un radio de más de seis kilómetros impedía de forma tajante la escolarización de la población. Las familias vivían al pie de las tierras que cultivaban y que les daban de comer, en el mejor de los casos en caserío propio o bien en las casas en las que servían como criados o bajo el régimen de aparcería. Los niños constituían mano de obra muy necesaria, sobre todo en determinadas épocas del año, período de siembra y de cosecha.

En la Tabla 2 se recoge la distribución de estas escuelas en 59 pueblos del concejo. En términos generales se observa que el tipo de centro predominante es la escuela mixta, en 31 de los 59 pueblos que tienen escuela este es el tipo de centro que existe, mientras que en 28 de los 59 pueblos existen escuelas unitarias de niños y niñas.

La existencia de escuelas unitarias de niños y niñas frente a las escuelas mixtas parece deberse a un criterio demográfico puesto que los núcleos con escuelas mixtas en principio son más pequeños, y por lo tanto, el número de niños en edad escolar es menor, es el caso de Alienes, Argumoso, Artosa, Biescas, Brañarronda y un largo etcétera.

Estos pueblos en la actualidad se encuentran muy despoblados, y los vecinos que permanecen suelen tener una edad avanzada. En menos de un siglo la realidad se ha tornado totalmente opuesta, mientras que a principios del siglo XX era perentorio el levantamiento de edificios escolares, ya a finales del mismo la mayoría de las escuelas rurales habían cerrado por falta de matrícula debido al brutal éxodo poblacional sufrido en esta zona.

**TABLA 2.** Estadística de escuelas en Valdés, 1958-1963

	<b>Gra- duada de niñas</b>	<b>Gradua- da de niños</b>	<b>Mixta</b>	<b>Párvu- los</b>	<b>Unita- ria de niñas</b>	<b>Unita- ria de niños</b>	<b>Total</b>
Aguera-Paredes							2
Alienes							1
Almuña							2
Arcallana							2
Argumoso							1
Artosa							1
Ayones							2
Barcia							3
Biescas							1
Brañarronda							1
Briefes							2
Buseco							1
Busto							2
Cadavedo							3
Cadollo							1
Canedo-Otur							1
Canero							3
Carboniella							2
Carcedo							2
Caroyas							2
Castañedo							2

Colinas							
Concernoso							
El Otero							2
El Vallín							
Ferrera							
La Mafalla							
La Montaña							2
La Mortera							2
Lago							
Leiriella							
Lendepeña							
Luarca							5
Menudero							
Moanes							2
Muñas							2
Otur							3
Paladeporre							
Paredes							2
Pereda							
Piñeros							
Pontigón							2
Quintana							2
Ranón							
Rellón de Merás							2
San Feliz							
San Pelayo de Tahona							
Santiago							2
Setienes							2
Sexmo							
Silvamayor							
Siñeriz							
Tablizo							
Telares							
Trevias							3
Valtravieso							
Villademos							
Villagermonde							
Villanueva-Cortina							2
<b>Total</b>	<b> </b>	<b> </b>	<b>31</b>	<b>6</b>	<b>28</b>	<b>28</b>	<b>95</b>

Fuente: Elaboración propia a partir de la Estadística de Enseñanza Primaria del Instituto Nacional de Estadística. Caja I 1292 del Archivo Histórico de Asturias.

Si analizamos el número de escuelas por parroquia se observa que en las de Canero, La Montaña y Santiago el número es de diez o más. En todos los pueblos que eran cabecera parroquial y, por lo tanto, conformaban el núcleo más poblado de la misma, existían escuelas unitarias de niños y de niñas. Esto es una regla excepto en la parroquia de Alienes, de poca extensión y muy baja densidad de población.

Si bien en el apartado 2.4 se profundiza en el conocimiento de las escuelas de párvulos, es posible adelantar que las seis unidades de párvulos se encuentran localizadas en las parroquias más pobladas, con una actividad económica importante y situadas en la rasa costera, a excepción de Trevías, que se enclava más hacia el interior del concejo. Lógicamente el grado de “modernización” de estos lugares es muy superior al de las parroquias más alejadas de la capital, plenamente sumergidas en una economía agraria próxima al autoabastecimiento en muchos sentidos.

TABLA 3. Estadística de escuelas por parroquia en el concejo de Luarca, 1958-1963

	Graduada de niñas	Graduada de niños	Mixta	Párvulos	Unitaria de niñas	Unitaria de niños	Total
Alienes			2				2
Arcallana			5		1	1	7
Ayones			1		1	1	3
Barcia			2	1	1	1	5
Cadavedo			1	1	1	1	4
Canero			1	1	4	4	10
Carcedo					2	2	4
Castañedo			1		1	1	3
La Montaña			8		2	2	12
Luarca	1	1		1	2	2	7
Muñas			2		2	2	6
Otur			1	1	1	1	4
Paredes			1		3	3	7
Santiago			2		4	4	10
Trevías			4	1	3	3	
<b>Total</b>	<b>1</b>	<b>1</b>	<b>31</b>	<b>6</b>	<b>28</b>	<b>28</b>	<b>95</b>

Fuente: Elaboración propia a partir de la Estadística de Enseñanza Primaria elaborada por el Instituto Nacional de Estadística. Caja 11292 del Archivo Histórico de Asturias.

El concejo de Valdés destaca por poseer el mayor número de brañas de toda Asturias, se trata de pequeñas poblaciones ubicadas en zonas elevadas y con pronunciadas pendientes. Once de estas brañas disponían de escuela propia, Artosa, Argumoso, El Vallín, Cadollo, Buseco, Brañaronda, Siñeriz, Silvamayor, Lendepeña, Leiriella y Lago.

La tipología arquitectónica es diversa en la medida en que responde a diferentes realidades, épocas y medios económicos. Así, vemos que los mejores edificios se concentran en Luarca y en pueblos pudientes como Arcallana, que cuenta con una construcción señorial en dos plantas. Respecto a la edificación de una gran parte de las escuelas unitarias construidas en el primer tercio del siglo XX, cabe señalar que presentan un diseño prácticamente idéntico, compuesto por un cuerpo central, que normalmente acogía la biblioteca, y dos alas, una destinada a niños y otra a niñas, dotadas de amplios ventanales a ambos lados (tres o cuatro dependiendo del tamaño de la misma) dando un conjunto alargado que proporcionaba mucha luminosidad. Es el caso de las escuelas de Muñás, Carcedo, Ferrera, Brieves, Villanueva, Silvamayor, Barcia, Cadavedo, Otur o Paredes.

Solamente en Luarca existía una escuela graduada de niñas y otra de niños, la Agrupación Escolar Mixta del Campo de la Feria, “con sus secciones en dos o más edificios” según la Estadística de la Enseñanza Primaria de 1965.



Foto de la Escuela de Muñás (Archivo de Eduardo Domínguez Sánchez-Pico)



Foto de la Escuela de Arcallana (Autora Trinidad Suárez Cano)

### 2.3. Sobre el origen de las escuelas valdesanas

En este apartado se trata de profundizar en el nacimiento de algunas de las escuelas del concejo tomando como hilo conductor principal la información procedente de la revisión de los fondos documentales disponibles en el Ayuntamiento de Valdés.

En términos generales, cabe destacar la aportación de los emigrantes para construir escuelas (Lombardero Rico, 2013). Participaban tanto los más favorecidos por la fortuna como aquellos más modestos, que a través de colectas enviaban lo que buenamente podían. Así, un elevado número de escuelas se levantaron con el dinero de estos emigrantes, especialmente antes del advenimiento de la Segunda República, momento en el que el Estado tomó conciencia de la importancia de la educación y estableció un plan de construcción de escuelas sin precedentes.

A modo de ejemplo, es necesario dejar constancia de la labor de Ramón Asenjo, emigrado a Argentina y a su regreso alcalde de Luarca durante varios años. Ramón Asenjo donó a título personal dinero y, además promovió desde la Alcaldía la construcción de varias escuelas

(Arcallana, Ayones, Barcia, Cadavedo, Luarca, Muñás, Canero, Paredes, Santiago, Trevías, La Carboniella, Carcedo y Otur, entre otras). Según el Informe de Primera Enseñanza de D. Benito Castrillo Sagredo (1926):

*En Ayones, D. Vicente Trelles dio 700 pesetas, D. Ramón Asenjo para varias escuelas una respetable cantidad; en Canero, D. Manuel Arango regaló el solar y dinero en total 4.000 pesetas, los emigrantes además dieron 1.687 pesetas, para la Biblioteca, y 1000 para una máquina de escribir; en Carcedo se recibieron de Cuba 13.000 pesetas; en Luarca parte del material y una máquina de coser, son de donación particular, la última de D. Francisco Gamoneda; en Muñás, los americanos dieron 9.000 pesetas, en Otur el Excelentísimo D. Ángel García, Viceconsul de España en Rosario de Santa Fe y su hermano D. Ramón aportaron 12.000; a Paredes llegaron de América 10.000, el solar lo regaló Dña. Germana González; en Santiago el solar lo donaron el popular banquero D. Vicente Trelles y el Sr. Marqués de Ferrera; en Trevías la suscripción popular produjo 19.000 pesetas, el Estado dio 5.000: tiene otro donativo para máquina de escribir y biblioteca circulante; en La Carboniella la suscripción alcanzó 7.500. En el citado informe de D. Benito Castrillo Sagredo (1926, p. 52-56) se señala: “en otras escuelas allí construidas no intervino para nada la acción del Estado”.*

Cabe destacar la creación de la escuela de Rellón de Merás por su peculiaridad, ya que fue levantada en 1913 por Manuel García Merás Sánchez, emigrante a Cuba, que cuenta además con dos viviendas adosadas para los maestros.

Otro caso análogo sucedió con la escuela de Lendepeña, inaugurada en 1927, construida a expensas de Nicolás Gayo, vecino del pueblo también emigrado a Cuba. Además de levantar el edificio escolar también llevó a cabo el camino que une Lendepeña a Arcallana, cabecera de la parroquia, y la capilla.

El testimonio de Valentín García Nido, nacido en Lendepeña en 1928, es bastante elocuente:

*Empecé a ir a la escuela del pueblo con 8 años. Tenía que caminar unos dos kilómetros por una subida muy pronunciada para llegar a la escuela, y los niños y niñas que venían de los pueblos de Arquillina, Cerezal (concejo de Salas) y Los Baos recorrían varios kilómetros todos los días para acudir a la escuela mi pueblo, lo que se volvía penoso cuando el frío y la lluvia arreciaban. La mitad de los días no venían.*

Buceando en el origen de la escuela graduada de Luarca, una de las más antiguas de la región, nos encontramos con que en 1906 el Ayuntamiento saca a subasta el proyecto para la construcción de unas escuelas públicas que él mismo tenía con anterioridad en la “Vega del Campo”, cuya edificación salió adelante con auxilio de los emigrantes. En 1910 dejan de ser escuelas municipales y se convierten en graduadas, que es como se las conocerá en lo sucesivo, mediante un Real Decreto, después de la Guerra Civil toman el nombre de “José Antonio” y en la actualidad “Padre Galo”.

En el ámbito rural, es necesario incidir en el hecho poco estudiado del papel protagonista que tuvieron los vecinos en la materialización del acceso a la enseñanza de los niños y niñas, que hasta el momento era muy deficiente en todos los aspectos. Fueron los propios vecinos quienes sufragaron parte de los gastos de construcción de los edificios, e incluso donaron los solares donde se habrían de levantar los mismos, y solicitaban a los ayuntamientos la contratación de un maestro/maestra, cuyo alojamiento también solía correr a cargo del vecindario.

De los casos encontrados cabe destacar la instancia que realizan los vecinos de los pueblos de Ore, Pontigón, Llaneces, Landedo y Bustiello al Sr. Alcalde-Presidente del Ayuntamiento y Junta Local de Primera Enseñanza de Luarca en la que se solicita la creación de una Escuela Nacional mixta en Pontigón (véase en el apéndice la transcripción de la instancia fechada en 1927). Además, en dicho documento se recoge el hecho de que las Escuelas Unitarias de Carcedo se crearon previamente siguiendo este mismo procedimiento que a continuación se expone, y todo parece indicar que fue el proceso usual en la construcción de la mayoría de las escuelas del concejo.

Los vecinos del pueblo de Piñeros de La Montaña sostenían una escuela desde por lo menos el año 1905, con la ayuda de una pequeña subvención del municipio consistente en 150 pesetas anuales, habiendo construido el vecindario exclusivamente por su cuenta el propio edificio. En el año 1931, en vista de los acuerdos del Ministerio de Instrucción Pública de levantar un considerable número de escuelas, se solicita la creación de un grupo de escuela mixta al que puedan acudir también los niños y niñas de Siñeriz y Degollado, pueblos que no contaban con escuela y que se encuentran a corta distancia. Una vez creada la escuela de Piñeros, los vecinos de Siñeriz también se movilizaron para crear una escuela en este pueblo.

Los vecinos de los pueblos de La Mortera, Mones, Espiniella Alta, Espiniella Baja y Candanosa, que forman el distrito escolar de La Mortera, en 1932, viendo que se había anunciado la subasta de las construcciones de los edificios escolares de Brieves, Silvamayor, Villagermonde, Siñeriz, Otero, Ferrera y Villanueva piden a la corporación municipal que incluya la construcción de una escuela en La Mortera ofreciendo el terreno para el solar y la cantidad en metálico de 3.480 pesetas. En Tablizo en 1933, cuyo distrito escolar abarca Tablizo, Gallinero y Busmarzo, la escuela funciona en un edificio de propiedad particular, arrendado por el Ayuntamiento, mientras se construye otro más adecuado.

En el municipio de Valdés se construyeron unas 32 escuelas entre el 28 de abril de 1931 y el 5 de octubre de 1934, periodo en el que fue alcalde Luis Ochoa de Albornoz (véase en el anexo la transcripción fechada el 8 de junio de 1932), lo que equivale a decir que más de un tercio de las escuelas del concejo fueron levantadas en este breve intervalo de tiempo.

## **2.4. Las escuelas de párvulos en valdés**

El afán benéfico propició el nacimiento de escuelas destinadas a recoger a niños de corta edad, procedentes de capas sociales desfavorecidas, e impulsado por la iniciativa privada (Molero Pintado, 1999). La revolución industrial supuso el primer gran trasvase de población desde las zonas rurales a las urbanas y la incorporación de la mujer al mercado de trabajo. Es en este contexto en el que surgen las primeras escuelas de párvulos, aunque todavía con un marcado carácter asistencial (Colmenar Orzaes, 1991).

Tal y como se recoge en el apartado anterior Valdés contaba con seis escuelas de párvulos: Luarca, Barcia, Otur, Canero, Cadavedo y Trevías, aunque según un documento hallado en el Archivo Municipal de Valdés, en 1932 se puso en funcionamiento otra escuela de párvulos en el casco de Luarca, mientras que la que se contabiliza en las estadísticas de los años sesenta sería la ubicada en el campo de la feria.

Estas localidades que contaban con escuela de párvulos eran las más cercanas a la villa de Luarca, con excepción de Trevías, que se sitúa en el interior del concejo. Si nos preguntásemos por las razones de su origen, además del número de habitantes y el polo de atracción que suponía Luarca como capital del concejo, veríamos que estas poblaciones tenían una actividad económica no estrictamente primaria como sí sucedía en

la mayor parte de los núcleos valdesanos, eminentemente agrícolas y ganaderos<sup>6</sup>.

En Luarca funcionó a partir de 1932 la escuela de párvulos del casco, creada por Orden de 26 de noviembre de 1932 (Gaceta del 1º de diciembre), en un edificio de propiedad particular, en el que con anterioridad se ubicaba la escuela de Artes y Oficios, arrendado por el Ayuntamiento y en el que se habían ejecutado las reformas pertinentes. Se puso en marcha con carácter provisional, mientras se construía otro más adecuado. Su creación responde a la creciente demanda de matrícula destinada a los niños más pequeños. Según la documentación manejada, se facilitaba a los alumnos el desayuno y la comida, pues se disponía de cocina destinada a tal fin.

Este último dato nos hace pensar que si bien la mayoría de los pueblos no contaba con una educación propiamente infantil plasmada en una escuela de párvulos ni ésta era obligatoria, en la práctica eran aceptados por los maestros si tenían voluntad de aprender y si sus padres así lo consideraban. La escuela de párvulos de Barcia se creó en 1933 dentro del proyecto de construcción de edificios escolares en todos los pueblos del concejo donde se ponía escuela, en palabras del alcalde republicano Luis Ochoa de Albornoz.

## Conclusiones

Hubo que esperar a los años veinte del pasado siglo para que se pudiese en funcionamiento una red escolar en el concejo atendida por maestros y maestras debidamente titulados y con unas instalaciones decorosas e higiénicas, pues hasta entonces se ponía escuela donde buenamente se podía, en hórreos, casas particulares o pajares, y se ocupaban de las mismas maestros carentes de la debida preparación.

Así pues, se deduce que las autoridades hasta ese momento no tenían como prioridad el establecimiento y consolidación de la educación primaria, y que tuvieron que ser espoleadas en cierta manera por los propios interesados para salir de su desidia.

La principal conclusión es que no es el Estado sino los vecinos de los pueblos los que demandan ayuda para abrir escuelas; contribuyen donando

---

<sup>(6)</sup> Para un análisis más detallado sobre el parvulismo en Valdés, véase Suárez y Suárez (2015).

el solar, a través de aportaciones económicas considerables para sus precarias economías y con su propio trabajo para la construcción y el acarreo de materiales donados por ellos mismos, principalmente piedra y madera.

Lo que se observa tras el análisis de la información recogida es que los padres, deseosos de que sus hijos tuviesen una vida mejor que la suya y de que se erradicase el analfabetismo en los pueblos, demandaron al Ayuntamiento de Luarca la creación de escuelas a lo largo y ancho del concejo, siendo el periodo que abarca la Segunda República el más fructífero en cuanto a levantamiento de edificios escolares, de la mano de su alcalde Luis Ochoa de Albornoz, que puso gran empeño en atender las demandas de los vecinos.

Otro aspecto a subrayar es la ingente labor que llevaron a cabo los emigrantes, siendo promotores exclusivos de varias escuelas y a través de aportaciones económicas que llegaban desde Cuba, Argentina, Méjico, Puerto Rico o Guatemala, contribuyendo a la construcción y al sostenimiento de numerosas escuelas. Cabe preguntarse qué hubiese ocurrido de no haber salido de sus hogares camino de las Américas tantos y tan generosos prohombres, que a fin de cuentas, con su labor de mecenazgo erradicaron el analfabetismo en sus añorados pueblos gracias al dinero que cruzó todo un océano.

Ciertamente, la creación de una escuela en un pueblo supuso fijar su población, mejorar el nivel cultural de sus habitantes y dotarlos de los instrumentos mínimos para encarar el tremendo cambio socioeconómico que experimentó España en los años sesenta. A día de hoy, las escuelas rurales de Valdés ya son pasado. Las últimas, Muñás y Concernoso, cerraron sus puertas en 2016 y 2017, respectivamente. Todas ellas han pasado a formar parte de ese nostálgico lugar en el que habita el olvido.

## Referencias Bibliográficas

- Bartolomé Cossío, M. (1985). *Una antología pedagógica*. Madrid: Ministerio de Educación.
- Bello, L. (1926). *Viaje por las escuelas de España*. Madrid: El Magisterio Español
- Colmenar Orzaes, C. (1991). Las escuelas de párvulos en España durante el siglo XIX: su desarrollo en la época de la Restauración. *Historia de la Educación*, vol. 10, 86-110.

- Egido Gálvez, I. (1994). La evolución de la enseñanza primaria en España: organización de la etapa y programas de estudio. *Tendencias pedagógicas*, nº 1, 75-86.
- Fernández Soria, J. M. y Agulló Díaz, M. C. (2004). *Una escuela rural republicana*. Valencia: Universitat de Valencia.
- Hernández Díaz, J. M. (2000). La escuela rural en la España del siglo XX. *Revista de Educación*, núm. extraordinario, 113-136.
- Lombardero Rico, Ch. (2013). *La buella de la emigración a América en el cantábrico occidental*. Gijón: CICEES.
- López Martín, R. (1997). La construcción y creación de escuelas en la España del primer tercio del siglo XX. *Historia de la educación*, nº 16, 65-90.
- López Martín, R. (1994). *Ideología y educación en la dictadura de Primo de Rivera. Escuelas y maestros*. Valencia: Universitat de Valencia.
- Martínez Sánchez, J.J. (2006). *La legislación escolar en la II República Española. Un referente de progreso y libertad en la escuela pública*. Available at: [http://martinez-carmona.es/libros/Legislacion\\_escolar\\_en\\_la\\_II\\_Republica\\_espanola.pdf](http://martinez-carmona.es/libros/Legislacion_escolar_en_la_II_Republica_espanola.pdf)
- Mato Díaz, A. (2010). Las escuelas y los maestros de primeras letras (siglo XIX). *Magister: revista miscelánea de investigación*, nº 23, 19-33.
- Mato Díaz, A. (1992). *La escuela primaria en Asturias (1923-1937). Los procesos de alfabetización y escolarización*. Oviedo: Dirección Provincial de Asturias, Ministerio de Educación y Ciencia.
- Molero Pintado, A. (1999). *Bases para una historia de la educación infantil en España. La figura de Eugenio Bartolomé y Mingo*. Madrid: Universidad de Alcalá.
- Núñez, C. E. (2005). Educación. En A. Carreras y X. Tafunell, X. (Ed.). *Estadísticas históricas de España Volumen I (155-244)*. Madrid: Fundación BBVA.
- Puelles Benítez, M. (1991). *Educación e ideología en la España contemporánea*. Barcelona: Editorial Labor.
- Soler Mata, J. (2008). El maestro y la “fisonomía propia” de la escuela rural. Una visión histórica de la escuela rural en Cataluña y España. En N. Llevot Calvet y J. Garreta Bochaca (Eds.) *Escuela rural y sociedad (11-41)*. Universitat de Lleida.
- Suárez, P., Suárez, T. (2015). Analfabetismo y creación de Escuelas en el Municipio de Valdés (Asturias). Available at: [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2700043](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2700043)

Tortella, G., Núñez, C.E. (2011). *El desarrollo de la España contemporánea. Historia económica de los siglos XIX y XX*. Madrid: Alianza Editorial.

**Información de contacto:** Patricia Suárez. Universidad de Oviedo, Facultad de Economía y Empresa, Departamento de Economía. Avenida del Cristo s/n. E-mail: suarezcpatricia@uniovi.es

## Apéndices

### **1. Transcripción de la instancia al Sr Alcalde- Presidente del Ayuntamiento y Junta Local de Primera Enseñanza de Luarca (Pontigón, Marzo 1º de 1927)**

Los que escriben, mayores de edad, Celadores Municipales de los pueblos de Ore, Pontigón y Llaneces de la parroquia de Carcedo, y de los de Landedo y Bustiello de la parroquia de Ayones, a V. y a las corporaciones de su digna presidencia exponen:

– Que en los referidos pueblos de Ore, Pontigón y Llaneces, con un censo de población aproximado de cuatrocientos habitantes, existen unos sesenta y cinco entre niños y niñas de edad escolar que carecen de la instrucción adecuada por no contar en la actualidad con otros medios de educación que una escuela particular subvencionada por el Ayuntamiento, instalada en pésimas condiciones (en un hórreo) careciendo del material necesario y de todo método pedagógico, teniendo por lo tanto que recorrer los niños distancias de 3, 4 y 5 kilómetros, por mal camino, hasta la Escuela Nacional más próxima, que es la de Carcedo, en la que, por otra parte, ya se cuenta con una matrícula escolar excesiva.

– Que en los también referidos pueblos de Landedo y de Bustiello, con un censo de población de cien habitantes y unos veinte y seis entre niños y niñas de edad escolar, no existen otros medios de educación que la Escuela Nacional mixta de Ayones, parroquia a la que pertenecen, distante unos diez kilómetros, también por malos caminos, siéndoles imposible durante el invierno y en época de lluvias recorrer esa distancia a los niños para asistir a la escuela.

– Que en vista de las grandes dificultades enumeradas, del deber ineludible que tenemos los padres de familia de velar por la instrucción de los hijos, cumpliendo así con una de nuestras obligaciones para con ellos y para con la Patria, cuya prosperidad y grandeza, a base de una mayor cultura, anhelamos; animados por las dignas Autoridades del Concejo y muy especialmente por el Sr. Delegado Gubernativo de este distrito, en reunión de todos los vecinos de estos pueblos, verificada en el Pontigón el día veintidós de Agosto de mil novecientos veinticuatro, quedó constituida una Junta patrocinadora de la creación de un grupo escolar en Pontigón o sus inmediaciones, por ser el punto céntrico de todos pueblos.

– Que en dicha reunión, según consta en la copia del acta que adjuntamos, se acordaron en principio las aportaciones de todos los vecinos ascendiendo las cantidades en metálico suscriptas a la suma de 3.334 pesetas que unidas a unas 1.200 y reunidas por suscripción en Cuba, dan un total de 4.534 pesetas, esperando que dichas suscripciones habrán de aumentar cuando se vea que la realización del proyecto es ya un hecho definitivo.

– Que en una reunión posterior verificada también en Pontigón a principios del mes de febrero último y convocada al efecto por el Sr. Cura regente de Carcedo D. Inocencio F. Villamil y por el concejal D. Antonio Castrillón García, a indicaciones de los mismos, se acordó por unanimidad ratificar el ofrecimiento anterior y ampliar el compromiso de las aportaciones del vecindario en la forma siguiente:

1º Ceder gratis al Ayuntamiento el solar para la edificación indicando el sitio donde ésta habrá de emplazarse, cosa que ya está resuelta con los actuales propietarios del terreno.

2º Comprometerse por prestación personal obligatoria al acarreo de todos los materiales necesarios para la construcción de la obra, como son: piedra, barro, arena, maderas, etc desde los lugares desde donde se extraigan y también de los demás materiales que se importen: cemento, cal y losa, desde el lugar donde sean descargados por los camiones u otros medios de transporte. Los vecinos que no posean carros apropiados a tal objeto se comprometen a abonar el importe de la prestación personal que se le señale por el celador respectivo.

– Y que, tanto las cantidades en metálico ya suscriptas como las que en lo sucesivo pudieran suscribirse, serán entregadas tan pronto como sea necesario, bien a la Comisión patrocinadora o bien al Ayuntamiento

según la Alcaldía disponga, siguiendo en esto el mismo procedimiento cuando la construcción de las Escuelas de Carcedo, obligándose todos al pago de dichas cantidades que podrán ser cobradas incluso por el procedimiento de apremio si ello fuese necesario.

– Por tanto, y a virtud de lo expuesto solicitamos primeramente la creación de una Escuela Nacional mixta en Pontigón, de ser posible regentada por Maestra, a cuyo efecto contamos en dicho pueblo con dos casas alquilables que creemos reúnen las debidas condiciones: una, de la que es apoderado el Sr. Cura párroco de Arcallana y está dispuesto a alquilar al Ayuntamiento para tal objeto; y otra de D. Facundo García, vecino de la localidad, el cual se ofrece espontáneamente, además, a facilitar casa –habitación gratis para la Maestra.

– Y por último, contando con las aportaciones del vecindario a cuyo cumplimiento se obliga con las firmas de todos puestas a continuación en pliegos aparte, que el Ayuntamiento de su digna presidencia, dando una vez más pruebas de su amor a la Enseñanza, acuerde incluir en el proyecto de las primeras construcciones de grupos escolares del concejo la de una escuela unitaria en el mencionado pueblo de Pontigón, parroquia de Carcedo, para que en ella puedan recibir instrucción los niños y niñas de los cinco pueblos interesados.

Es justicia que esperamos alcanzar de V. cuya vida guarde Dios muchos años. Pontigón, Marzo 1º de 1927.

El Celador de Ore José Pérez, El Celador de Pontigón Higinio Gómez  
El Celador de Llaneces Diego Álvarez, El Celador de Landedo Jesús  
García

El Celador de Bustiello José Fernández, El Alcalde de Barrio Florentino  
Fernández

## **2. Transcripción de la instancia que realiza D. Luis Ochoa de Albornoz (Sr Alcalde- Presidente del Ayuntamiento) al Ilustrísimo Director General de Primera Enseñanza de la provincia de Oviedo (Luarca, 8 de Junio de 1932)**

D. Luis Ochoa de Albornoz, Alcalde- Presidente del Ayuntamiento de Luarca, acogiéndome a lo dispuesto en el Real Decreto de 10 de Julio de 1928.

A V. I. acude solicitando la subvención de DIEZ Y OCHO MIL PESETAS para la construcción directa por el Municipio de un edificio de nueva planta con destino a dos escuelas nacionales unitarias, una para niños y

otra para niñas, en el pueblo de Villanueva, de este concejo, a cuyo efecto se acompaña la oportuna certificación de los acuerdos municipales, el proyecto redactado por el Arquitecto D. José Ramón del Valle y el croquis o plano de emplazamiento para el expresado edificio escolar.- Luarca ocho de Junio de mil novecientos treinta y dos.

# Coaching a docentes y rendimiento académico: PTA en Colombia<sup>1</sup>

## Linking coaches and student achievement: PTA in Colombia

DOI: 10.4438/1988-592X-RE-2017-381-385

Ximena Dueñas Herrera  
Ana Bolena Escobar Escobar  
Silvana Godoy Mateus  
Jorge Leonardo Duarte Rodríguez

*Instituto Colombiano para la Evaluación de la Educación*

### Resumen

El Programa para la Transformación de la Calidad Educativa *Todos a Aprender* (PTA) es un programa de Desarrollo Profesional Situado o *Coaching*, implementado en Colombia por el Ministerio de Educación Nacional desde el año 2012. El PTA está dirigido a los profesores de educación Básica Primaria de los colegios oficiales del país que tienen gran cantidad de estudiantes con dificultades en su proceso de aprendizaje. La estrategia del programa incluye la selección y capacitación de formadores y tutores encargados de apoyar a los docentes en el desarrollo de sus clases, y la provisión de materiales educativos para docentes y estudiantes. La revisión de la literatura existente permite concluir que el impacto del programa en el desempeño académico de los estudiantes no ha sido evaluado; por lo cual, se utilizan los resultados obtenidos durante cuatro años consecutivos en la prueba Saber 3°, 5° y 9°, aplicada por el Instituto Colombiano para la Evaluación de la Educación (Icfes) a nivel de grado, área y jornada de los establecimientos, para intentar medir el efecto del PTA mediante la estimación de un modelo lineal y un modelo Logit con efectos fijos a nivel de sede. Los resultados muestran que, en la mayoría de grados y áreas, un año adicional en el PTA se correlaciona positivamente con el desempeño de los

---

<sup>(1)</sup> Agradecemos a Diana Oquendo Victoria por la edición del documento.

estudiantes. En particular, el efecto del programa es más robusto en el grado tercero y el área de matemáticas para las sedes del nivel socioeconómico más bajo. Aunque estos resultados deben leerse con algunas precauciones, proporcionan información que puede ser útil para la toma de decisiones relacionadas con la ampliación y continuidad del programa.

*Palabras clave:* Desarrollo Profesional Situado, Coaching, Tutores, Evaluación de Impacto, Desempeño Académico.

### **Abstract**

The Transformation of Educational Quality Program *Todos a Aprender* (PTA) is a Situated Professional Development or *Coaching* program, implemented by the Ministry of National Education since 2012. The program is aimed at the country's primary education teachers of public schools that have a large quantity of students with difficulties in their learning process. The program's strategy includes the selection and training of trainers and tutors to support teachers in the development of their classes, and the provision of educational materials to teachers and students. The review of the existing literature shows that the impact of the program on students' academic performance has not been evaluated so far. Hence, with data from the standardized test Saber 3°, 5° y 9°, applied during four consecutive years by the Colombian Institute for the Evaluation of Education (Icfes) at grade, area and school shift, we try to measure the effect of PTA through the estimation of a linear model and a Logit model with fixed effects on the school performance. The results show that, in most grades and areas, an additional year in PTA correlates positively with student's performance. Moreover, the effect is more robust in the third grade and math area in schools with a lower socioeconomic level. Although these results must be read with caution; the results may be useful for decision-making regarding the extension and continuity of the program.

*Keywords:* Situated Professional Development, Coaching, Tutors, Impact Evaluation, Academic Performance

## **Introducción**

En educación, el Desarrollo Profesional Situado (DPS) o *Coaching* es una estrategia implementada para apoyar a los docentes en el desarrollo de sus clases mediante la asignación de tutores que les provean apoyo planificado y continuo en su proceso de enseñanza (Department

of Education and Early Childhood Development, 2010). Los tutores desarrollan actividades de planificación conjunta, análisis de práctica docente, registran información sobre el proceso de acompañamiento e implementan modalidades de trabajo que respondan a los problemas detectados (Vezub y Alliaud, 2012).

El efecto del coaching en la práctica docente puede variar dependiendo del tiempo que el tutor apoya al profesor, de la obligatoriedad o voluntariedad de la participación del docente, y del contenido y contexto del acompañamiento. El DPS puede incidir en el ámbito pedagógico y didáctico al promover la ejecución de estrategias efectivas de enseñanza y proveer conocimiento sobre el origen y tratamiento de las dificultades que enfrentan los estudiantes en su proceso de aprendizaje (Duque, Celis, Díaz y Gómez, 2014).

En Colombia, el Ministerio de Educación Nacional (MEN) ejecuta desde 2012 el programa *Todos a Aprender* (PTA), que provee DPS voluntario a docentes de Básica Primaria de establecimientos oficiales caracterizados por tener muchos estudiantes con rendimiento académico bajo. En particular, el programa asigna tutores para que acompañen a los profesores de los colegios seleccionados, y ofrece materiales de trabajo y lineamientos sobre los procedimientos a realizar y los objetivos a alcanzar (MEN, 2016).

El PTA es relevante en el contexto colombiano por dos razones principales. Primero, por las altas brechas en el acceso a la educación de calidad y en las diferencias en el rendimiento académico de los estudiantes según sus características socioeconómicas (Icfes, 2016). Y segundo, porque existe una baja capacidad de restructuración de la planta docente en los colegios oficiales y es inusual que los profesores sean despedidos por su bajo desempeño (OCDE, 2009a). En este sentido, aunque el PTA no es una solución completa, sí es una estrategia que busca reducir estos problemas desde el corto plazo.

Aunque hay numerosos programas de coaching para profesores en el mundo y diferentes aproximaciones a la evaluación de su impacto, la revisión de la literatura existente muestra que el impacto del PTA en Colombia no ha sido evaluado. Por ello, este documento investiga el efecto del PTA en las competencias de matemáticas y lenguaje de los estudiantes focalizados por el programa.

Para el desarrollo de la investigación se emplean algunas características contextuales de los estudiantes y los establecimientos educativos (EE),

los resultados de las pruebas *Saber* aplicadas por el Instituto Colombiano para la Evaluación de la Educación (Icfes)<sup>2</sup> y la pertenencia al PTA de todos los colegios del país para un periodo de cinco años, contados a partir de 2012, año en el que empezó el programa. Este panel de datos permite usar modelos de efectos fijos con el fin de controlar por características no observables invariantes en el tiempo, lo que permite disminuir las amenazas a la validez estadística asociadas a la naturaleza observacional de los datos.

Este documento se distribuye de la siguiente manera: Tras la introducción, se presentan experiencias internacionales de programas de DPS y los métodos empleados para evaluar su impacto. Posteriormente, se describe el funcionamiento del PTA y el mecanismo de selección de los colegios beneficiados. Seguidamente, se describen los datos y la metodología empleada en el ejercicio empírico. Por último, se presentan los resultados y una sección de conclusiones.

## **Desarrollo Profesional Situado: Experiencias Internacionales**

En los últimos años se ha destacado la importancia del desarrollo integral de los docentes, que se entiende como la formación de profesionales con las competencias necesarias para desarrollar actividades que fomenten el aprendizaje. En otras palabras, lo fundamental no es poseer el conocimiento académico, sino la capacidad de transmitirlo y enseñar a usarlo a través de las experiencias (OECD, 2010). La importancia del adecuado desarrollo de este proceso en el aprendizaje de los estudiantes ha dado lugar a la implementación de programas de DPS. Este concepto define una relación colaborativa que ayuda a los docentes a obtener aprendizaje profesional en su lugar de trabajo, posibilidad que les permite mejorar su práctica de enseñanza en el aula (Rodríguez, Esteban, Aranda, Blanchard, Domínguez, González, Romero, Sanz, Mamposo, Vitón y Messina, 2011).

Los programas de DPS pueden ser clasificados entre obligatorios o voluntarios y de acuerdo al tipo de instrucción que ofrecen, donde encontramos formación inicial de docentes o apoyo a profesores de

---

<sup>(2)</sup> Organismo oficial encargado de realizar pruebas estandarizadas por competencias para distintos niveles de la educación en Colombia.

bajo desempeño (Mourshed, Chijioke y Barber, 2010; Solaz-Portolés y López, 2008). En algunos países de Europa, Asia y Oceanía existen programas obligatorios para profesores principiantes que son instruidos y acompañados hasta que adquieren la capacidad para entrenar a otros docentes (Darling-Hammond, Chung y Andree, 2010). En Estados Unidos y Canadá, además de los programas obligatorios de acompañamiento a profesores principiantes (Darling-Hammond, et al., 2010; Flanagan, 2006; Lockwood, Sloan y Marsh, 2010), existen programas dirigidos a docentes de colegios con bajo rendimiento (Mourshed, et al., 2010; UNESCO, 2014).

En Latinoamérica la implementación de programas de tutoría para docentes principiantes empezó en 2000 (Vezub y Alliaud, 2012). En Argentina el Instituto Nacional de Formación Docente coordina el acompañamiento de instituciones de educación superior a docentes principiantes (Congreso de la Nación Argentina, 2007). En Chile un programa de DPS selecciona mentores formados en universidades focalizadas para que ayuden a los docentes mediante la planificación conjunta, la observación, el análisis recíproco de clases, la retroalimentación y la reflexión sobre las prácticas de enseñanza (Boerr, 2010). Finalmente, en Uruguay se establecen “parejas pedagógicas” de profesores que brindan apoyo mutuo a partir de la observación y el análisis de la práctica cotidiana (Rodríguez y Grilli, 2012).

Con el objetivo de mejorar la práctica docente, en México y Ecuador, además del acompañamiento a profesores principiantes, existen programas de DPS para profesores que presenten un nivel de desempeño insuficiente en la evaluación docente (Congreso General de los Estados Unidos Mexicanos, 2013; Vezub y Alliaud, 2012). Con el mismo objetivo, en el Estado Minas Gerais de Brasil se ejecuta un programa que contempla cuatro visitas semanales a colegios de bajo desempeño, en las que se capacita a los rectores para que desempeñen el rol de “coaches” de los docentes (Mourshed, et al., 2010).

## **Evaluaciones**

Numerosas investigaciones han evaluado el impacto de los programas de DPS en la calidad educativa (Cumsille y Fiszbein, 2015). Dichos estudios

emplean métodos de análisis cualitativo, cuantitativo o mixto<sup>3</sup> y se han concentrado en la medición del efecto de este tipo de programas en las creencias y prácticas de los maestros, y en el rendimiento académico de los estudiantes. Las investigaciones que realizan análisis cualitativo usualmente emplean los datos de planificación de las lecciones del docente y el tutor, el seguimiento a las sesiones y los cuestionarios o entrevistas con el tutor, el docente o el director de la institución educativa (Ensminger, Kallemeyn, Rempert, Wade y Polanin, 2015; Flanagan, 2006; Meher, Ummulbanin y Mursaleen, 2007; Neuberger, 2012).

Los estudios cuantitativos enfocados en el impacto de los programas sobre las prácticas de enseñanza docente han empleado múltiples metodologías. Villar y Strong (2007) elaboran un análisis costo beneficio de un programa de mentoría para profesores principiantes calculando la tasa de retorno y encuentran que, después de cinco años, una inversión en el programa produce un retorno positivo para la sociedad, el distrito escolar, los maestros, los estudiantes y el Estado. Bartlett (2012), Veenman, De Laat y Staring (1998) y Veenman y Denessen (2001) evalúan las habilidades de los profesores antes y después del acompañamiento y encuentran que estos programas mejoran las habilidades de enseñanza, fortaleciendo la autonomía de los docentes, la retroalimentación y el estímulo a la reflexión. Además, encuentran que la experiencia del mentor, la duración de la tutoría y la definición clara de los objetivos del programa inciden positivamente en el rendimiento, la confianza y las aspiraciones de los profesores (Veenman, et al., 1998; Veenman y Denessen, 2001).

En el análisis cuantitativo del efecto de los programas de coaching sobre el desempeño de los estudiantes, los investigadores estiman diferentes modelos. Matsumura, Garnier, Junker, Resnick y DiPrima (2009) emplean un modelo lineal jerárquico y concluyen que, aunque el programa aumenta la calidad de la enseñanza de los maestros, el rendimiento de los estudiantes no presenta un cambio significativo. Lockwood, Sloan y Marsh (2010) estiman un modelo de efectos fijos y encuentran una relación positiva con el desempeño de los estudiantes. Biancarosa, Bryk y Dexter (2010) aplican un modelo de valor agregado y evidencian que un año adicional de pertenencia a este tipo de programas aumenta el aprendizaje de los estudiantes. Por su parte, Yoshikawa, Snow, Barata, Gómez, Leyva,

---

<sup>3</sup> Cappella, Jackson, Kim, Bilal, Hollad y Atkins (2016) y Sumner (2011) realizan análisis mixto.

Treviño, Weiland, y Moreno (2015) usan un modelo multinivel, a partir del cual identifican impactos positivos en la organización de la clase y en el apoyo emocional e instruccional a los docentes, y afirman que no hay impacto en las habilidades socioemocionales de los estudiantes ni en su aprendizaje en lectoescritura.

## **Programa Todos a Aprender**

El PTA es un programa de DPS destinado a profesores de establecimientos oficiales que enseñan en el nivel educativo de Básica Primaria (grados de transición a quinto) en Colombia. Este programa intenta formar a los docentes en aspectos disciplinares y didácticos, mediante tutores que los acompañan en el desarrollo de sus lecciones y trabajan con ellos en sesiones adicionales, que buscan mejorar las condiciones de aprendizaje de los estudiantes (MEN, 2015). Cuantitativamente, el programa busca que más del 25 por ciento de los estudiantes de los colegios focalizados asciendan de nivel de desempeño, al menos en las áreas de lenguaje y matemáticas evaluadas en la prueba Saber 3°, 5° y 9° (MEN, 2013).

El PTA involucra cuatro componentes claves para un adecuado acompañamiento a los profesores. El componente pedagógico proporciona referentes curriculares, materiales educativos de lenguaje y matemáticas para profesores y estudiantes, y alternativas de evaluación de aprendizajes. El componente de condiciones básicas brinda soporte en alimentación, transporte e infraestructura física y tecnológica para garantizar la permanencia de los estudiantes en el sistema educativo. El componente de gestión educativa posibilita la evaluación constante del programa. Y el componente de comunicación, movilización y compromiso promueve el involucramiento de toda la comunidad educativa (MEN, 2011).

La estrategia de acompañamiento a los profesores consta de formadores y tutores. Los formadores son contratados por el MEN y los tutores son seleccionados por las entidades territoriales certificadas (ETC). Aunque los tutores pertenecen a la ETC, el presupuesto para el pago de tutores y formadores es del MEN. Los procesos de formación del PTA funcionan en forma de cascada: los formadores instruyen a los tutores, quienes, a su vez forman a los profesores de acuerdo con un cronograma y una agenda definida (MEN, 2013). Desde el año 2017 los formadores serán escogidos por universidades y acompañarán directamente a los profesores.

El PTA está dirigido a colegios que están ubicados en regiones donde las brechas educativas son notorias y que tienen estudiantes con dificultades en el logro de aprendizajes (MEN, 2015). Por ello, el MEN, la firma consultora McKinsey & Co y las ETC eligieron articuladamente los establecimientos beneficiarios del programa<sup>4</sup>. Inicialmente, el MEN seleccionó los 1.616 EE más rezagados en calidad educativa, con base en los resultados de las pruebas *Saber*, la tasa de repitencia (mayor a 4.6%), la tasa de deserción (mayor a 3.8%) y las variaciones negativas en la matrícula de los EE presentadas entre 2009 y 2010. Adicionalmente, consideró un componente de regionalización y criterios relacionados con indicadores socioeconómicos y de vulnerabilidad de los EE.

La consultora McKinsey & Co consideró la matrícula y la facilidad de acceso al establecimiento para incrementar el impacto del PTA en cobertura y costo eficiencia, con lo que seleccionó 1.400 EE adicionales. En total, del proceso de selección del MEN y la consultora McKinsey & Co resultaron 3.016 EE, de los cuales solo 2.378 fueron focalizados como beneficiarios del PTA. Sin embargo, antes de implementar el programa, algunas ETC le solicitaron al MEN la inclusión de algunos de sus EE. En total, 1.933 EE hacen parte de este grupo, de los cuales el 90% está en la categoría de desempeño mínimo o bajo/prioritario.

Aunque los establecimientos que iniciaron como beneficiarios del PTA continúan en la focalización, desde 2016 se definió que la duración de la intervención será de tres años. Este límite se estableció porque, desde la creación del programa, se espera que el acompañamiento posibilite la construcción de ideas sobre la clase de prácticas que los profesores pueden lograr y sostener por sí mismos (MEN, 2013). Actualmente el PTA beneficia a 2.400.000 niños y 92.000 docentes en más de 4.300 instituciones de 878 municipios del país (MEN, 2016).

## Metodología

### Representatividad de las bases de datos

En esta investigación se utilizan dos bases de datos. La primera contiene los resultados obtenidos, entre 2012 y 2016, por los grados tercero y quinto de cada sede o establecimiento del país en la prueba *Saber 3°*,

---

<sup>(4)</sup> La información proviene de documentos del programa y diversas conversaciones y reuniones con miembros del equipo del MEN, del PTA y de la consultora McKinsey.

5° y 9°. Esta evaluación es aplicada anualmente por el Icfes a nivel de área y grado para cada sede o establecimiento del país. A nivel de establecimiento, el Icfes brinda un puntaje promedio y el porcentaje de estudiantes según cuatro niveles de desempeño (Insuficiente, Satisfactorio, Mínimo o Avanzado); sin embargo, a nivel de sede, no hay puntaje promedio y a nivel de estudiante no se publica ningún resultado. Además de la evaluación, el Icfes aplica a los estudiantes cuestionarios que recopilan información sobre sus condiciones socioeconómicas y contextuales, las características de su colegio y la comunidad donde este se ubica. En particular, es de interés el Índice de Nivel Socioeconómico, que cataloga a los colegios en una de cuatro categorías socioeconómicas, donde el nivel socioeconómico (NSE) 1 es la más baja. El cálculo del NSE tiene en cuenta el nivel educativo de los padres o acudientes del estudiante y la infraestructura, servicios y posesiones de la vivienda.

La segunda base de datos fue suministrada por el MEN e indica el año de ingreso al programa de todas las sedes que fueron focalizadas entre 2012 y 2016. Una característica importante para este trabajo es que, aunque los preseleccionados del PTA son EE, en la práctica la focalización está dirigida a las sedes<sup>5</sup>. La focalización del PTA es anual y se ha realizado siempre a principio del año calendario, que coincide con al año escolar de los colegios oficiales. A diferencia de los datos del Icfes, que ofrecen una calificación para cada grado, área, jornada y sede, la implementación del PTA es para toda la sede. Esta característica permite hacer análisis separados para cada grado, área y jornada. Dado que cada sede tiene un identificador único mediante el cual es posible relacionar las bases de datos del MEN y el Icfes, la base de datos resultante empleada en la investigación es longitudinal a nivel de área (matemáticas o lenguaje), grado (tercero o quinto) y jornada de la sede (mañana, tarde o completa), y contiene información sobre la evaluación, las características contextuales y la pertenencia al programa desde 2012 hasta 2016. Además, debido a que el ingreso al PTA está restringido a EE oficiales, la base de datos resultante solo contiene sedes del sector oficial.

La Tabla 1 muestra la cantidad de sedes registradas en las bases de datos del Icfes y el MEN, así como las observaciones para las cuales fue posible cruzar información de ambas bases. Como lo evidencian los datos, en todos los años de estudio, el porcentaje de sedes beneficiarias del PTA para el que fue posible vincular información de la prueba *Saber*,

---

<sup>5</sup> Una sede es una de varias plantas educativas que pertenecen a un mismo EE. Entre sedes existen diferencias en términos de ubicación, planta docente y población de estudiantes atendidos. Debido a costos en el acceso, cuando un EE con muchas sedes es seleccionado, solamente algunas de sus sedes son focalizadas.

en una o más jornadas, varía entre el 40 y el 63 por ciento para grado tercero, entre el 38 y el 59 por ciento para grado quinto, y entre el 79 y el 100 por ciento para ambos grados.

Entre las razones por las cuales no fue posible vincular información de la prueba a algunas de las sedes participantes del PTA se encuentran: sedes sin grados tercero y quinto, cambios o problemas en el código de identificación de sede, incumplimiento de los requisitos que habilitan la publicación de resultados, entrega de las hojas de respuesta de la prueba sin distinción de sedes (en estos casos solo es posible generar resultados a nivel de establecimiento), entre otros. A pesar de que el número de observaciones perdidas en el proceso de emparejamiento de las bases no es despreciable, la cantidad de sedes que cuentan con información sobre el año de ingreso al programa y los resultados de la prueba *Saber*, en una o más jornadas, sigue siendo considerablemente alto (entre 4.470 y 15.342).

TABLA I. Número de sedes en las bases empleadas por año y jornada

	Sedes con resultados en pruebas Saber		PTA	Cruce	
	Tercero	Quinto		Tercero	Quinto
<b>Mañana</b>					
<b>2012</b>	15.402	14.130	9.517	3.681	3.399
<b>2013</b>	12.170	11.376	23.166	6.129	5.749
<b>2014</b>	14.481	13.541	24.045	7.774	7.202
<b>2015</b>	15.185	14.227	25.466	8.552	7.958
<b>2016</b>	19.080	17.867	24.408	10.622	9.864
<b>Tarde</b>					
<b>2012</b>	2.348	2.253	9.517	743	714
<b>2013</b>	2.027	1.944	23.166	1.026	1.012
<b>2014</b>	2.625	2.504	24.045	1.415	1.380
<b>2015</b>	2.592	2.482	25.466	1.479	1.422
<b>2016</b>	2.815	2.678	24.408	1.569	1.506
<b>Completa</b>					
<b>2012</b>	6.979	6.622	9.517	379	357
<b>2013</b>	6.104	5.897	23.166	2.194	2.149
<b>2014</b>	6.777	6.638	24.045	2.552	2.510
<b>2015</b>	6.237	6.071	25.466	2.498	2.476
<b>2016</b>	8.176	7.934	24.408	3.151	3.086

\*Nota: La implementación del PTA se realiza por sedes y no por jornadas, por lo que no es posible establecer qué jornadas son focalizadas por el programa. Una misma sede puede tener varias jornadas.

Fuente: Elaboración propia

## Características de los datos

La Tabla 2 muestra la frecuencia de sedes oficiales que pertenecen y no pertenecen al PTA por NSE. Las sedes reportadas como ND no están catalogadas en un NSE debido a dificultades al momento de llenar el cuestionario. La tabla muestra el crecimiento de la focalización del PTA en los años del estudio: mientras que, en 2012, el programa atendía casi el 20 por ciento de las sedes del sector oficial, en 2016 ya acompañaba a más de la mitad. Tanto la focalización del PTA como su crecimiento en el tiempo se concentran en las sedes de los NSE más bajos y sobre todo en el NSE 2. En el año 2016, las sedes del NSE 1 representan casi el 45 por ciento de las sedes focalizadas por el programa; y las de los NSE 1 y 2, el 98 por ciento.

**TABLA 2.** Número de sedes oficiales con y sin PTA, por NSE y año

	<b>Año</b>	<b>No</b>	<b>Sí</b>
NSE 1	2012	9.732	2.205
	2013	4.716	3.919
	2014	5.320	4.828
	2015	5.060	5.577
	2016	7.250	7.039
NSE 2	2012	10.095	2.511
	2013	5.558	5.375
	2014	6.129	6.782
	2015	5.716	6.799
	2016	7.075	8.393
NSE 3	2012	966	109
	2013	802	209
	2014	951	280
	2015	916	275
	2016	988	299

NSE 4	2012	16	0
	2013	20	0
	2014	16	0
	2015	20	0
	2016	19	0
N.D.	2012	228	13
	2013	163	11
	2014	79	11
	2015	80	6
	2016	81	7
Total	2012	21.037	4.838
	2013	11.259	9.514
	2014	12.495	11.901
	2015	11.792	12.657
	2016	15.413	15.738

Fuente: Elaboración propia

Aunque el programa está enfocado en los establecimientos de bajo NSE, la Tabla 2 muestra que la focalización no está estrictamente restringida a este tipo de establecimientos. Esta circunstancia se da por diferentes razones. La aceptación y reconocimiento del programa ha hecho que este crezca en el tiempo y que más establecimientos quieran participar. Las ETC son intermediarias en el proceso y pueden tener incentivos para postular ciertos establecimientos. El PTA no es obligatorio, así que los EE preseleccionados pueden decidir no participar.

Como la selección de los beneficiarios del programa no es aleatoria, no es sencillo estimar el impacto del PTA en el desempeño académico de los estudiantes. Sin embargo, las circunstancias anteriormente mencionadas sobre la selección de las sedes participantes proporcionan un escenario favorable para el análisis (o estrategia de identificación), pues permiten comparar sedes relativamente similares con y sin PTA.

La Tabla 3 muestra el porcentaje de estudiantes de sedes oficiales que, en promedio, se encuentran en los niveles de desempeño Insuficiente y Mínimo, con y sin PTA, para grado tercero en el área de matemáticas en la jornada de la mañana. En estos datos resaltan tres características principales: hay muchas más sedes en los NSE bajos que en los altos; a medida que disminuye el NSE, aumenta el porcentaje de estudiantes

en los dos niveles de desempeño más bajos, es decir, disminuye el desempeño de los estudiantes; y tanto las sedes beneficiarias del PTA como las que no, parecen mejorar en el tiempo.

**TABLA 3.** Porcentaje de estudiantes de sedes oficiales, de grado tercero, área de matemáticas, jornada mañana, en los niveles de desempeño Insuficiente y Mínimo con y sin PTA, por NSE y año<sup>6</sup>

		Sin PTA			Con PTA		
		N*	Media	Desv. Est.	N	Media	Desv. Est.
NSE 1	2012	5.187	49,88	43,27	1.626	63,75	40,39
	2013	2.271	43,17	42,57	2.338	65,18	39,81
	2014	2.779	49,12	42,77	2.984	62,14	40,10
	2015	2.716	35,81	41,86	3.611	54,28	42,19
	2016	3.739	36,74	42,87	4.474	43,57	42,02
NSE 2	2012	4.791	49,67	36,60	1.737	62,64	33,47
	2013	2.473	47,91	36,25	3.099	56,28	35,76
	2014	2.677	46,33	35,15	4.008	53,61	34,88
	2015	2.639	40,77	36,52	4.032	51,79	36,19
	2016	3.116	42,91	37,42	4.906	44,45	36,96
NSE 3	2012	628	39,39	20,52	67	43,06	21,77
	2013	521	34,90	19,15	131	38,34	22,48
	2014	596	32,79	19,28	191	37,81	20,68
	2015	580	35,28	19,87	186	39,20	22,39
	2016	577	42,40	10,84	191	43,57	23,31
NSE 4	2012	7	16,15	15,49			
	2013	10	28,53	22,74			
	2014	7	8,27	4,23			
	2015	9	13,79	10,04			
	2016	8	28,65	14,94			
N.D.	2012	125	46,99	44,92	9	69,53	36,37
	2013	81	42,41	47,07	4	56,81	41,72
	2014	51	36,60	45,59	9	46,30	44,70
	2015	58	21,33	39,44	5	40,00	54,77
	2016	56	19,35	38,67	6	16,67	40,82
<b>Total</b>		<b>35.702</b>			<b>33.614</b>		

\* Número de sedes en las que se ubica el porcentaje de estudiantes reportado.

Fuente: Elaboración propia

<sup>6)</sup> La tendencia de los resultados es similar para las demás combinaciones entre grados, áreas y jornadas.

## Métodos

Para intentar estimar el impacto del PTA en el desempeño de los estudiantes, en un contexto de no aleatoriedad en la selección del tratamiento, se emplean modelos de efectos fijos a nivel de sede para cada jornada, grado y área, de la siguiente manera:

$$y_{jt} = f(\beta_0 + \beta_1 + PTA_{jt} + \beta_2 X_{jt} + \delta_i + \mu_{it}) \quad (1)$$

Donde,  $f(.)$  es una función lineal o Logit, en la que  $j = (1, 2, \dots, N)$  es la jornada, área y grado de una sede y  $t = (2012, 2013, 2014, 2015$  y  $2016)$  es el año para el que se disponen resultados en las pruebas *Saber*.  $y_{jt}$  es el resultado de la evaluación de los estudiantes de la sede  $j$  en el año  $t$ .  $PTA_{jt} = (0, 1, 2, 3, 4, 5)$  es una variable que cuenta el número de años de permanencia de una sede en el PTA<sup>7</sup>, de esta manera, una sede que nunca ha estado en el programa tiene el valor de cero y una que está desde el año 2012 tiene el valor de uno en 2012, de dos en 2013, hasta el valor de cinco en 2016. Además,  $X_{jt}$  recopila las características de la sede, que varían en el tiempo<sup>8</sup>,  $\delta_i$  las características observables y no observables invariantes en el tiempo y  $\mu_{it}$  es el término de error. Finalmente,  $\beta_1$  es el coeficiente de interés que, dada una buena especificación del modelo, muestra el efecto que tiene un año adicional del programa en el desempeño promedio de los estudiantes de la sede.

La elección un modelo de efectos fijos se debe a la correlación existente entre las características no observables invariantes en el tiempo y la pertenencia al PTA, explicada por la no aleatoriedad en la selección de los beneficiarios del programa. Por ejemplo: características observables correlacionadas con el PTA son el NSE de los estudiantes y la ubicación de la sede; y características no observables pueden ser la disposición del rector para querer participar en el programa y la habilidad intrínseca de profesores y estudiantes. Si no se tuvieran en cuenta los efectos fijos a nivel de sede (correlacionados con la pertenencia al programa), estos quedarían incluidos en el término de error y  $\beta_1$  estaría sesgado.

<sup>(7)</sup> La especificación de la variable PTA permite analizar la intensidad del efecto y disponer de múltiples controles y tratamientos con características similares.

<sup>(8)</sup> La mayoría de características demográficas y socioeconómicas que recopila el Icfes varían poco en el tiempo; por lo cual,  $X_{jt}$  es únicamente el número de evaluados en la jornada, grado y área de la sede.

Dado que la implementación del PTA tiene lugar en las sedes<sup>9</sup>, se emplea el porcentaje de estudiantes por niveles de desempeño para construir dos variables de respuesta. La primera es la suma del porcentaje de estudiantes en los niveles de desempeño Insuficiente y Mínimo, que puede considerarse una variable apropiada porque el objetivo del PTA es reducir la cantidad de niños en los niveles de desempeño más bajos. La segunda variable dependiente es una dummy que indica si la sede ha aumentado de un año a otro el porcentaje de niños en los niveles más bajos, por lo que toma el valor de uno si la diferencia de un año a otro entre la suma del porcentaje de estudiantes en los niveles Insuficiente y Mínimo es mayor que cero, y, en caso contrario, toma el valor de cero.

Con la variable en niveles, se estima un modelo lineal que intenta calcular el efecto que tiene un año adicional de pertenencia al PTA en la disminución del porcentaje de estudiantes en los niveles de desempeño Insuficiente y Mínimo. Con la variable dummy se estima un modelo Logit que muestra la dirección del efecto que tiene un año adicional del PTA en la probabilidad de aumentar el porcentaje de estudiantes en los niveles de desempeño más bajos. En el modelo Logit quedan tres observaciones por sede debido a la variable dependiente en diferencias.

Con el objetivo de analizar sedes que tienen una probabilidad similar de haber sido seleccionadas como beneficiarias del programa, las observaciones de los modelos solo incluyen las sedes que pertenecen a los NSE 1 y 2. Para cada combinación de áreas, grados y jornadas, se hacen dos especificaciones: la primera incluye todas las sedes oficiales del grupo, y la segunda considera únicamente aquellas que han participado en el PTA por lo menos un año. La razón de hacer esta exclusión es que las sedes que no han participado en el programa en ninguno de los cinco años podrían tener características no observables sistemáticamente diferentes a las que sí han participado. La Tabla 4 muestra la cantidad de sedes oficiales de los NSE 1 y 2 para las dos especificaciones descritas anteriormente y puede entenderse como un filtro de la Tabla 1, debido a que ambas especificaciones son obtenidas seleccionando las sedes oficiales de NSE 1 y 2 de la base del Icfes y de aquella que contiene las sedes para las cuales fue posible cruzar información de las bases de datos del Icfes y el MEN.

---

<sup>9)</sup> En la prueba Saber 3°, 5° y 9° el Icfes reporta el puntaje promedio únicamente a nivel de establecimiento.

**TABLA 4.** Número de observaciones en cada especificación, por año y jornada

	Sedes oficiales con resultados publicados en pruebas Saber NSE I y 2		Cruce NSE I y 2	
	Tercero	Quinto	Tercero	Quinto
<b>Mañana</b>				
<b>2012</b>	14.533	13.346	3.602	3.320
<b>2013</b>	11.389	10.650	5.993	5.616
<b>2014</b>	13.602	12.763	7.567	7.016
<b>2015</b>	14.325	13.495	8.356	7.783
<b>2016</b>	18.212	17.099	10.421	9.676
<b>Tarde</b>				
<b>2012</b>	1.792	1.714	686	665
<b>2013</b>	1.509	1.455	934	927
<b>2014</b>	1.977	1.904	1.283	1.259
<b>2015</b>	1.962	1.871	1.343	1.293
<b>2016</b>	2.146	2.053	1.424	1.373
<b>Completa</b>				
<b>2012</b>	6.850	6.498	376	355
<b>2013</b>	5.970	5.765	2.172	2.130
<b>2014</b>	6.652	6.516	2.536	2.497
<b>2015</b>	6.098	5.939	2.478	2.457
<b>2016</b>	7.987	7.746	3.123	3.059

Fuente: Elaboración propia

## Resultados

Como se dijo anteriormente, existe un número no despreciable de sedes beneficiarias del PTA para las cuales no fue posible vincular información de la prueba Saber 3°, 5° y 9°, como consecuencia, los resultados presentados a continuación son válidos para las sedes con respecto a las cuales tenemos información del año de ingreso al programa y de los resultados en la prueba.

## Efecto del PTA sobre el porcentaje de estudiantes en Mínimo e Insuficiente

En el modelo lineal, el coeficiente de interés representa el efecto marginal promedio de un año adicional en el programa sobre el porcentaje de estudiantes en los niveles de desempeño Insuficiente y Mínimo. Este porcentaje, condicionado por las características de las sedes invariantes en el tiempo y el número de estudiantes evaluados en cada año, permite comparar el efecto entre los grados, áreas y jornadas consideradas. Los resultados (Tabla 5) indican que un año adicional en el PTA se correlaciona con una disminución promedio del porcentaje de estudiantes en los niveles de desempeño más bajos, es decir, el PTA se asocia con una mejora promedio en el desempeño de los estudiantes en los grados y áreas consideradas, donde el grado tercero en el área matemáticas registra el mayor impacto. Aunque la forma funcional del modelo lineal puede ser inadecuada, debido a que la variable dependiente es un porcentaje, se emplea porque los resultados de las pruebas *Saber* tienen una distribución normal y la cantidad de sedes sin estudiantes en los dos niveles de desempeño más bajos o con todos los estudiantes en estos niveles es mínima.

TABLA 5. Coeficiente del PTA en el modelo lineal por grado, área y jornada

	LENGUAJE				MATEMÁTICAS			
	Tercero		Quinto		Tercero		Quinto	
	Solo PTA	Oficiales						
<b>Jornada Completa</b>								
Años PTA	<b>-4,44***</b>	<b>-3,51***</b>	<b>-2,01***</b>	<b>-1,30***</b>	<b>-4,28***</b>	<b>-3,33***</b>	<b>-2,35***</b>	<b>-2,71***</b>
	(-9,75)	(-10,09)	(-5,10)	(-4,33)	(-9,41)	(-9,85)	(-5,77)	(-8,84)
N	9.510	29.807	9.806	30.187	9.541	29.771	9.753	29.918
<b>Jornada Mañana</b>								
Años PTA	<b>-4,50***</b>	<b>-4,23***</b>	<b>-2,44***</b>	<b>-2,05***</b>	<b>-5,23***</b>	<b>-4,67***</b>	<b>-3,02***</b>	<b>-3,05***</b>
	(-22,65)	(-24,48)	(-14,22)	(-13,61)	(-25,05)	(-26,13)	(-17,54)	(-20,27)
N	32.622	65.052	31.837	63.635	32.815	65.203	31.675	63.233
<b>Jornada Tarde</b>								
Años PTA	<b>-2,65***</b>	<b>-2,84***</b>	<b>-2,13***</b>	<b>-2,02***</b>	<b>-3,26***</b>	<b>-3,14***</b>	<b>-2,29***</b>	<b>-2,36***</b>
	(-9,04)	(-10,88)	(-10,02)	(-10,48)	(-10,25)	(-10,95)	(-10,30)	(-11,66)
N	5.480	9.129	5.495	8.965	5.548	9.208	5.488	8.957

Estadístico t en paréntesis

\* p<0,05 \*\* p<0,01 \*\*\* p<0,001

Fuente: Elaboración propia

## Efecto del PTA sobre la probabilidad de aumentar el porcentaje de estudiantes en Mínimo e Insuficiente

Con el objetivo de conocer la dirección que tiene el efecto de un año adicional de PTA en la probabilidad de aumentar el porcentaje de estudiantes en los niveles de desempeño más bajos, se estimó un modelo Logit. En este caso, la variable de respuesta tiene forma de valor agregado, de modo que permite comparar el cambio de un año a otro entre las sedes participantes y no participantes del PTA. Las mediciones en forma de valor agregado tienen reconocidas ventajas sobre las medidas en niveles, debido a que permiten mejorar la condición observacional de las pruebas estandarizadas, al comparar los cambios y no solo las condiciones de salida (Downes y Vindurampulle, 2007; Drury y Doran, 2003; OCDE, 2008; Willms, 2009). Dado que, en promedio, las sedes focalizadas en el PTA tienen un desempeño académico más bajo que las no focalizadas, esta puede ser una buena especificación. La Tabla 6 muestra los resultados del modelo Logit y evidencia que el efecto es negativo y significativo para todos los grados, áreas y jornadas.

En general, los modelos usados permiten concluir que, para los grados tercero y quinto en el área de matemáticas y para el grado tercero en el área de lenguaje, el PTA se asocia con una mejora en los resultados de los estudiantes. Además, para todos los grados y áreas de las jornadas completa y mañana, el PTA se asocia con una disminución en la probabilidad de no mejorar.

### Robustez

La estimación del modelo lineal y del modelo Logit se realizó junto con algunas verificaciones de robustez. En los anexos 1 y 2 se presenta el modelo lineal estimado de forma independiente para el NSE 1 y 2, respectivamente. Este ejercicio, que permite comparar controles y tratamientos con características similares, también se realizó con el modelo Logit, donde se encontró el resultado esperado en todos los grados, áreas y jornadas.

Los anexos evidencian que los resultados de la Tabla 5 se mantienen para el caso del NSE 1 y que la magnitud del efecto es más grande. Este hallazgo puede ser interesante para los tomadores de decisiones

sobre políticas públicas, debido a que sugiere que los recursos son más eficientes si se destinan a la población más pobre. Sin embargo, es importante recordar que el desempeño promedio de las sedes de NSE 1, tanto sin como con PTA, es considerablemente más bajo que las de los otros NSE, por lo que la necesidad de mejora es mucho mayor.

**TABLA 6.** Coeficiente del PTA en el modelo Logit por grado, área y jornada

	LENGUAJE				MATEMÁTICAS			
	Tercero		Quinto		Tercero		Quinto	
	Solo PTA	Oficiales						
<b>Jornada Completa</b>								
<b>Años PTA</b>	<b>-0,23***</b>	<b>-0,18***</b>	<b>-0,10*</b>	<b>-0,06*</b>	<b>-0,25***</b>	<b>-0,17***</b>	<b>-0,13**</b>	<b>-0,18***</b>
	(-6,86)	(-6,99)	(-2,46)	(-2,00)	(-7,72)	(-7,00)	(-3,24)	(-5,84)
<b>N</b>	3.961	14.705	3.081	12.049	4.274	15.510	3.060	11.096
<b>Jornada Mañana</b>								
<b>Años PTA</b>	<b>-0,27***</b>	<b>-0,24***</b>	<b>-0,18***</b>	<b>-0,14***</b>	<b>-0,34***</b>	<b>-0,29***</b>	<b>-0,26***</b>	<b>-0,26***</b>
	(-12,25)	(-13,00)	(-6,01)	(-5,77)	(-16,58)	(-16,63)	(-8,78)	(-10,64)
<b>N</b>	9.146	23.225	5.775	16.312	10.507	25.636	5.505	15.172
<b>Jornada Tarde</b>								
<b>Años PTA</b>	-0,18	<b>-0,24*</b>	-0,44	-0,40	<b>-0,30**</b>	<b>-0,23*</b>	<b>-0,62*</b>	<b>-0,67*</b>
	(-1,44)	(-2,19)	(-1,54)	(-1,66)	(-2,68)	(-2,34)	(-2,06)	(-2,40)
<b>N</b>	271	514	60	102	329	622	55	82

Estadístico t en paréntesis

\* p<0,05 \*\* p<0,01 \*\*\* p<0,001

Fuente: Elaboración propia

## Conclusiones

### Hallazgos más significativos y recomendaciones de política pública

Esta investigación estima el impacto del PTA sobre las habilidades promedio de los estudiantes de grados tercero y quinto, medidas a través de la prueba Saber 3°, 5° y 9°. Los resultados evidencian que, para la mayoría de grados y áreas estudiadas, hay una correlación parcial y positiva entre el desempeño de los estudiantes y un año adicional de pertenencia al programa. Sin embargo, es importante aclarar que no se analizaron los canales mediante los cuales opera el programa. Es decir,

se encuentra que en promedio los establecimientos focalizados por el PTA están mejorando moderadamente más que los no focalizados, pero se desconoce el motivo que explica esta diferencia.

Adicionalmente, los resultados evidencian que el impacto es más fuerte en los establecimientos que llevan más tiempo en la focalización y en los que pertenecen al NSE más bajo, lo que sugiere que el programa tiene rendimientos crecientes y es más eficiente en los establecimientos de más bajo NSE. Estos hallazgos son útiles para que los encargados de las políticas públicas tomen decisiones respecto de continuar y ampliar el programa, mejorar las estrategias de acompañamiento, sofisticar el mecanismo de selección de los establecimientos y priorizar la focalización a los establecimientos de más bajo NSE.

## **Limitaciones y posibilidades de mejora**

Es importante reconocer que estas estimaciones pueden estar sesgadas debido a características no observables que varíen en el tiempo y debido a características comunes entre las observaciones para las cuales se cuenta con información sobre la pertenencia al programa y los resultados de la prueba. Aun así, teniendo en cuenta el alto número de sedes beneficiarias del programa para las cuales fue posible vincular los resultados de la prueba y bajo el supuesto de que no ha habido cambios considerables en el tiempo relacionados con el PTA, que afecten sistemáticamente el desempeño de los estudiantes, estos resultados sugieren que el programa está generando una modesta pero perceptible mejora en el desempeño de los estudiantes. Otra de las razones por las cuales estos resultados deben leerse con precaución es la característica observacional de los datos.

La aproximación empírica usada puede ser mejorada y complementada de diferentes maneras. El MEN está consolidando una base de datos de registros administrativos que comprende el número de visitas al aula, sesiones de los tutores con los formadores y profesores, y materiales entregados. Además, está clasificando a los tutores cualitativamente con base en algunos criterios de desempeño. La interacción de estas variables con la pertenencia al programa puede posibilitar el entendimiento del mecanismo por el cual el PTA se correlaciona positivamente con el desempeño de los estudiantes y permitir la definición de recomendaciones específicas sobre la forma de intervenir.

Finalmente, consideramos que quedan algunas preguntas importantes sin resolver sobre las posibles implicaciones del PTA en el proceso de aprendizaje de los estudiantes. Futuras investigaciones pueden evaluar si los beneficios del programa superan sus costos e indagar los mecanismos por los cuales la participación en el programa puede conducir a mejoras en el rendimiento académico de los estudiantes.

## Referencias bibliográficas

- Bartlett, J. (2012). *A model role - Evaluation of the Mosaic mentoring programmes*. London: Demos.
- Barrera, F., Maldonado, D. y Rodríguez, C. (2012). Calidad de la educación básica y media en Colombia: Diagnóstico y propuestas. *Serie documentos de trabajo*, 126.
- Biancarosa, G., Bryk, A. y Dexter, E. (2010). Assessing the value-added effects of literacy collaborative professional development on student learning. *The elementary school journal*, 111(1), 7 – 34.
- Boerr, I. (2010). *Acompañar los primeros pasos de los docentes*. Chile: Santillana.
- Cappella, E., Jackson, D., Kim, H., Bilal, C., Hollad, S. y Atkins, M. (2016). Implementation of teacher consultation and coaching in urban schools: A mixed method study. *School Mental Health*, 8(2), 222 – 237.
- Cid, A., Pérez, A. y Sarmiento, J. (2011). La tutoría en el Practicum. Revisión de la literatura. *Revista de Educación*, 354, 127 – 154.
- Congreso de la Nación Argentina. (2007). Ley de Educación Nacional 26.206. Disponible en: [http://www.me.gov.ar/doc\\_pdf/ley\\_de\\_educ\\_nac.pdf](http://www.me.gov.ar/doc_pdf/ley_de_educ_nac.pdf)
- Congreso General de los Estados Unidos Mexicanos. (2013). Ley General del Servicio Profesional Docente. Disponible en: [http://www.dof.gob.mx/nota\\_detalle.php?codigo=5313843&fecha=11/09/2013](http://www.dof.gob.mx/nota_detalle.php?codigo=5313843&fecha=11/09/2013)
- Cumsille, B. y Fiszbein, A. (2015). *Construyendo políticas docentes efectivas: Pautas para la acción*. Comisión para la Educación de Calidad para todos: Documento de antecedentes.
- Darling-Hammond, L., Chung, R. y Andree, A. (2010). *How High-Achieving Countries Develop Great Teachers*. Stanford, CA: Stanford Center for Opportunity Policy in Education.

- Department of Education and Early Childhood Development. (2010). *Coaching Teachers in Effective Instruction*. Melbourne: Student Learning Division Office for Government School Education.
- Downes, D. y Vindurampulle, O. (2007). *Value-added Measures for School Improvement*. Melbourne: Department of Education and Early Childhood Development.
- Duque, M., Celis, J., Diaz, B. y Gómez, M. (2014). Diez pilares para un programa de desarrollo profesional docente centrado en el aprendizaje de los estudiantes. *Revista Colombiana de Educación*, 67, 107 – 124.
- Drury, D. y Doran, H. (2003). The Value of Value-Added Analysis. *Policy Research Brief. National School Boards Association*, 3(1), 1 – 4.
- Ensminger, D., Kallemeyn, L., Rempert, T., Wade, J. y Polanin, M. (2015). Case study of an evaluation coaching model: Exploring the role of the evaluator. *Evaluation and Program Planning*, 49, 124 – 136.
- Flanagan, T. (2006). *The perceived effectiveness of a beginning teacher mentoring program in central Virginia*. (Tesis inédita de doctorado). Liberty University.
- García, S. (1999). El desarrollo profesional: análisis de un concepto complejo. *Revista de Educación*, 318, 175 – 187.
- Icfes. (2016). Informe Nacional Saber 3°, 5° y 9°. Resultados Nacionales 2009 – 2014.
- Lockwood, J., Sloan, J. y Marsh, J. (2010). Linking reading coaches and student achievement: Evidence from Florida middle schools. *Educational Evaluation and Policy Analysis*, 32(3), 372 – 388.
- Matsumura, L., Garnier, H., Junker, B., Resnick, L. y DiPrima, D. (2009). *The influence of content-focused coaching on reading comprehension instruction and Student Achievement*. SREE Conference abstract template.
- Meher, R., Ummulbanin, A. y Mursaleen, G. (2007). Teacher mentoring programme: A vehicle to support professional development for improving the quality of education in Districts of Sindh and Balochistan of Pakistan. *Quality in Education: Teaching and Leadership in Challenging Times*, 354 – 366.
- MEN. (1979). Decreto N°2.277 de 1979. Estatuto docente. Septiembre 14 de 1979.
- MEN. (2011). Programa para la transformación de la calidad educativa: Sustentos del programa.

- MEN. (2013). Orientaciones técnicas para la producción de secuencias didácticas para un Desarrollo Profesional Situado en las áreas de matemáticas y ciencias. Bogotá.
- MEN. (2015). *¿Qué es Todos a Aprender?* Obtenido de Colombia Aprende: <http://aprende.colombiaaprende.edu.co/es/pta/87535>
- MEN. (2016). *Cifras: Colombia Aprende*. Obtenido de Colombia Aprende: <http://aprende.colombiaaprende.edu.co/es/pta/87542>
- Mourshed, M., Chijioke, C. y Barber, M. (2010). *How the world's most improved school systems keep getting better*. London: McKinsey & Company.
- Neuberger, J. (2012). Benefits of a teacher and coach collaboration: A case study. *The Journal of Mathematical Behavior*, 31(2), 290 – 311.
- OCDE. (2008). *Measuring improvements in learning outcomes: Best practices to assess the value-added of schools*. Paris: OECD.
- OCDE. (2009 a). *Los docentes son importantes: atraer, formar y conservar a los docentes eficientes*. OECD Publishing.
- OCDE. (2009 b). *Evaluación y reconocimiento de la calidad de los docentes. Prácticas internacionales*. OECD Publishing.
- OECD. (2010). *OECD Reviews of Human Resource Management in Government: Brazil 2010: Federal Government*, OECD Publishing, Paris. <http://dx.doi.org/10.1787/9789264082229-en>
- Rodríguez, A., Esteban, R., Aranda, R., Blanchard, M., Domínguez, C., González, P., Romero, P., Sanz, E., Mampaso, A., Vitón, M. y Messina, C. (2011). *Coaching reflexivo entre iguales en el Practicum de la formación de maestros*. *Revista de Educación*, 355, 355 – 379.
- Rodríguez, E. y Grilli, J. (2012). *Transitar hacia la profesión docente. La pareja pedagógica como dispositivo en la formación inicial de profesores para la educación media en Uruguay*. III Congreso Internacional sobre profesorado principiante e inserción profesional a la docencia, Santiago de Chile.
- Sánchez, M. y Corte, F. (2015). La evaluación a la docencia: Algunas consecuencias para América Latina. *Revista Mexicana de Investigación Educativa*, 20(67), 1233 – 1253.
- Solaz-Portolés, J. y López, V. (2008). Types of knowledge and their relations to problem solving in science: directions for practice. *Educational Sciences Journal*, 6, 105-112.
- Sumner, K. (2011). *An explanatory mixed-methods study of instructional coaching practices and their relationship to student achievement*. (Tesis inédita de doctorado). Western Carolina University.

- UNESCO. (2014). *Temas críticos para formular nuevas políticas docentes en América Latina y el Caribe: el debate actual*. Santiago: OREALC/UNESCO Santiago.
- Veenman, S., De Laat, H. y Staring, C. (1998). Evaluation of a coaching programme for mentors of beginning teachers. *Journal of In-Service Education*, 24(3), 411 – 431.
- Veenman, S. y Denessen, E. (2001). The coaching of teachers: results of five training. *Educational Research and Evaluation*, 7(4), 385 – 417.
- Vezub, L. y Alliaud, A. (2012). *El acompañamiento pedagógico como estrategia de apoyo y desarrollo profesional de los docentes noveles: Aportes conceptuales y operativos para un programa de apoyo a los docentes principiantes de Uruguay*. OEI.
- Villar, A. y Strong, M. (2007). Is mentoring worth the money? A benefit-cost analysis and five-year rate of return of a comprehensive mentoring program for beginning teachers. *Educational Research Service*, 25(3), 1 – 17.
- Willms, D. (2009). *Value Added Models in Education*. Policy Brief. Canadian Research Institute for Social Policy.
- Yoshikawa, H., Snow, C., Barata, M., Gómez, C., Leyva, D., Treviño, E., Weiland, C. y Moreno, L. (2015). Experimental impacts of a teacher professional development program in Chile on Preschool Classroom quality and child outcomes. *Developmental Psychology*, 51(3), 309 – 322.

**Dirección de contacto:** Ximena Dueñas Herrera. Instituto Colombiano para la Evaluación de la Educación, ICFES. Bogotá. Carrera 4, n° 26<sup>a</sup>-42, Apartamento 317, Bogotá. E-mail: duenax@gmail.com

## Anexos

### ANEXO I. Modelo lineal solo para NSE I

	LENGUAJE				MATEMÁTICAS			
	Tercero		Quinto		Tercero		Quinto	
	Solo PTA	Oficiales						
<b>Jornada Completa</b>								
<b>Años PTA</b>	<b>-5,92***</b>	<b>-4,79***</b>	<b>-2,12***</b>	<b>-1,97***</b>	<b>-6,51***</b>	<b>-4,93***</b>	<b>-2,19***</b>	<b>-2,99***</b>
	(-7,43)	(-8,25)	(-3,24)	(-4,08)	(-8,55)	(-8,55)	(-3,21)	(-5,79)
<b>N</b>	3.879	12.197	4.023	12.394	3.881	12.146	3.993	12.240
<b>Jornada Mañana</b>								
<b>Años PTA</b>	<b>-5,46***</b>	<b>-4,92***</b>	<b>-2,99***</b>	<b>-2,61***</b>	<b>-7,18***</b>	<b>-6,19***</b>	<b>-3,69***</b>	<b>-3,69***</b>
	(-16,05)	(-16,86)	(-10,42)	(-10,55)	(-19,88)	(-20,46)	(-13,27)	(-15,51)
<b>N</b>	14.932	31.594	14.554	30.951	15.033	31.725	14.470	30.726
<b>Jornada Tarde</b>								
<b>Años PTA</b>	<b>-5,40***</b>	<b>-5,00***</b>	<b>-2,85***</b>	<b>-3,07***</b>	<b>-7,43***</b>	<b>-7,07***</b>	<b>-3,22***</b>	<b>-3,28***</b>
	(-4,85)	(-5,24)	(-4,30)	(-5,17)	(-6,85)	(-7,22)	(-4,45)	(-5,00)
<b>N</b>	599	765	628	784	618	787	625	780

Estadístico t en paréntesis

\* p<0,05 \*\* p<0,01 \*\*\* p<0,001

Fuente: Elaboración propia

### ANEXO II. Modelo lineal solo para NSE 2

	LENGUAJE				MATEMÁTICAS			
	Tercero		Quinto		Tercero		Quinto	
	Solo PTA	Oficiales						
<b>Jornada Completa</b>								
<b>Años PTA</b>	<b>-3,59***</b>	<b>-2,72***</b>	<b>-1,94***</b>	<b>-0,91*</b>	<b>-2,97***</b>	<b>-2,35***</b>	<b>-2,43***</b>	<b>-2,56***</b>
	(-6,50)	(-6,29)	(-3,94)	(-2,37)	(-5,26)	(-5,45)	(-4,81)	(-6,73)
<b>N</b>	5.631	17.610	5.783	17.793	5.660	17.625	5.760	17.678
<b>Jornada Mañana</b>								
<b>Años PTA</b>	<b>-3,81***</b>	<b>-3,72***</b>	<b>-2,04***</b>	<b>-1,64***</b>	<b>-3,85***</b>	<b>-3,56***</b>	<b>-2,54***</b>	<b>-2,58***</b>
	(-15,96)	(-17,72)	(-9,70)	(-8,75)	(-15,53)	(-16,51)	(-11,64)	(-13,35)
<b>N</b>	17.690	33.458	17.283	32.684	17.782	33.478	17.205	32.507
<b>Jornada Tarde</b>								
<b>Años PTA</b>	<b>-2,37***</b>	<b>-2,61***</b>	<b>-2,05***</b>	<b>-1,92***</b>	<b>-2,81***</b>	<b>-2,72***</b>	<b>-2,19***</b>	<b>-2,27***</b>
	(-7,84)	(-9,68)	(-9,14)	(-9,41)	(-8,50)	(-9,12)	(-9,36)	(-10,67)
<b>N</b>	4.881	8.364	4.867	8.181	4.930	8.421	4.863	8.177

\* p<0,05 \*\* p<0,01 \*\*\* p<0,001

Estadístico t en paréntesis

Fuente: Elaboración propia



# Estado de las prácticas científicas e investigación educativa. Posibles retos para la próxima década

## State of scientific practices and educational research. Potential challenges for the next decade

DOI: 10.4438/1988-592X-RE-2017-381-386

Ángeles Blanco-Blanco

Universidad Complutense de Madrid

### Resumen

El trabajo presenta una revisión del estado de las prácticas científicas actuales y de su impacto potencial en la calidad de la investigación educativa. Desde una concepción post-positivista de la investigación científica en educación, el problema se desarrolla en el contexto general del debate actual sobre la ciencia, sobre su fiabilidad, su robustez y su reproducibilidad. Teórica y conceptualmente el estudio adopta el enfoque propio de la meta-investigación. Desde el punto de vista metodológico, se lleva a cabo una revisión de la literatura sobre el tema que permita una reflexión fundamentada sobre el *status quo* de las prácticas científicas. Para ello se consideran algunos de los trabajos más relevantes y de mayor impacto publicados en los últimos años sobre meta-ciencia, en general y en el ámbito de las Ciencias de la Educación y del Comportamiento en particular. En primer lugar se caracteriza lo que se ha etiquetado como *crisis de la ciencia* y particularmente se trata de la presencia de sesgos y prácticas de investigación cuestionables en la investigación científica. A continuación se presentan algunos de los elementos correctores clave que se han propuesto para fortalecer y permitir un avance más efectivo de la ciencia. Estos incluyen: alternativas y “nuevos” énfasis en el análisis estadístico de datos científicos; el impulso renovado a la replicación y la reproducibilidad; y los nuevos modos de producción, difusión y evaluación de la investigación ligados a la ciencia abierta. El artículo se cierra con algunas reflexiones relativas a posibles retos para la investigación educativa española en la próxima década. Las conclusiones se organizan en torno a cuatro ejes: el desarrollo de estudios meta-investigativos;

la formación, información y sensibilización de los investigadores sobre prácticas de investigación cuestionables; la actualización de las políticas editoriales; y el papel de los financiadores y evaluadores de la producción científica.

*Palabras clave:* Prácticas de investigación cuestionables, Crisis de reproducibilidad, Ciencia abierta, Ecosistemas de la producción y la publicación científicas, Meta-investigación, Investigación educativa.

### **Abstract**

This paper presents a review of the state of current scientific practices and their potential impact on the quality of educational research. From a post positivist conception of scientific research in education, the matter is addressed in the general context of the current debate about science, its reliability, robustness and reproducibility. Theoretically and conceptually, the study adopts a meta-research approach. From a methodological perspective, a review of the literature on the subject is carried out that allows a reasoned reflection on the *status quo* of scientific practices. This is done by considering some of the most relevant papers published in recent years on meta-science in general, and in the field of Education and Behavioral Sciences. In the first place, the so-called *crisis of science* is characterized, and particularly the presence of biases and questionable research practices in scientific research. Below are some of the key corrective elements proposed to strengthen and enable a more effective advancement of the scientific enterprise. These include: alternative and “new” emphasis on the statistical analysis of scientific data; renewed impetus to replication and reproducibility; and new modes of production, dissemination and assessment of research associated with open science. The article closes with some reflections regarding possible challenges for Spanish educational research in the next decade. The conclusions are organized around four axes: the development of meta-research studies; training, information and awareness-raising of researchers about questionable research practices; updating editorial policies; and the role of funders and evaluators of scientific production.

*Key words:* Questionable research practices, Reproducibility crisis, Open Science, Scientific ecosystem, Meta-research, Educational research.

La investigación empírica en educación está llamada a informar las prácticas y las intervenciones educativas. Pero para poder contribuir de un modo efectivo a la toma de decisiones una condición necesaria, aunque ciertamente no suficiente, es que la investigación educativa sea

científicamente robusta, fiable y válida. En este estudio proponemos una reflexión sistemática sobre dicha condición. Presentamos una revisión de las prácticas científicas actuales para analizar su impacto potencial sobre la calidad y el rigor de la investigación educativa. El análisis se realiza desde una concepción post-positivista de la investigación científica en educación y en el contexto general del debate actual sobre la ciencia, sobre su fiabilidad, su robustez y su reproducibilidad. Desde el punto de vista teórico-conceptual el estudio adopta el enfoque propio de la meta-investigación. Desde el punto de vista metodológico, se lleva a cabo una revisión de la literatura sobre el tema que permita una reflexión fundamentada acerca del *status quo* de las prácticas científicas actuales. Para ello se consideran algunos de los trabajos más relevantes y de mayor impacto publicados en los últimos años sobre meta-ciencia, en general y en el ámbito de las Ciencias de la Educación y del Comportamiento en particular.

## Crisis de la ciencia y meta-investigación

En los últimos años se ha asistido en un amplia variedad de campos disciplinares a una fuerte crisis de confianza en lo que podríamos llamar el *status quo* de las prácticas científicas. Se ha hablado de *crisis de la ciencia* y, particularmente, de una *crisis de reproducibilidad* (Munafò et al., 2017). Un buen indicador del estado de cosas en este sentido lo representan las editoriales simultáneas publicadas por las revistas *Science* y *Nature* en 2014 bajo un elocuente título: *Journals unite for reproducibility* (Nature, 2014; McNutt, 2014). La editorial da cuenta de un acuerdo conjunto para adoptar nuevos principios y directrices en la publicación de investigación clínica. Pero lo esencial y que queremos destacar aquí es que se trataba así de reforzar las iniciativas que en los últimos años se venían tomando como respuesta a la crisis generada por las evidencias de falta de reproducibilidad, fraudes y malas prácticas que venían minando la confianza en la ciencia no sólo de los profesionales, sino de la opinión pública en general.

La preocupación y el estado de opinión en la comunidad científica puede ser ilustrado con algunos datos obtenidos en una encuesta reciente publicada en *Nature* (Baker, 2016). El 90% de los 1576 investigadores encuestados cree que efectivamente se asiste a una crisis

de reproducibilidad y el 52% la califican como significativa. Las causas más destacadas incluyen prácticas de investigación cuestionables (p.e. informe selectivo o *p-hacking*), la presión por publicar y cuestiones vinculadas al uso del aparato estadístico y el diseño de investigación. Esta última dimensión, la estadístico-metodológica, encabeza la lista de soluciones señaladas por los encuestados, particularmente lo referido a la mejora en la formación y en la supervisión estadístico-metodológica. De hecho se ha apuntado que la *crisis* podría estar en buena parte derivada del mal uso del apartado estadístico en la investigación científica (Goodman, 2016; Peng, 2015).

Las dificultades encontradas para reproducir la investigación publicada deben principalmente verse como un síntoma, como el resultado de un problema de fondo complejo vinculado a diversos factores que expresan el escenario y los parámetros bajo los cuales se hace y difunde la ciencia. Como recuerdan Makel y Plucker (2014), que un resultado no sea replicable no le convierte en falso. Y obviamente un resultado replicable no es por ello verdadero. Pero ciertamente la replicabilidad es una pieza angular del sistema de la ciencia y la discusión en torno a la misma parece haber puesto un foco de atención necesario sobre el rigor y la transparencia de las prácticas científicas<sup>1</sup>.

Donde posiblemente más fuerte y turbulenta ha sido la crisis en los últimos años es en el ámbito de la investigación en Psicología (Ledgerwood, 2014), un campo relativamente afín al educativo. La detección de fraudes a cargo de reputados investigadores, la publicación de dudosos y controvertidos descubrimientos y la evidencia de prácticas cuestionables extendió la sospecha de que quizá no todo funciona bien tras las elevadas tasas de éxito con la que los investigadores predicen los resultados que hallan en sus propios estudios. Así Simmons, Nelson & Simonsohn (2011) sugirieron que la hiper-precisión en la predicción (p.e. rechazo de la hipótesis nula) es probablemente debida a varios factores, que incluyen: recogida de datos hasta que se encuentra el resultado

---

<sup>(1)</sup> Efectivamente la replicación, entendida como la repetición de estudios previos a cargo de investigadores independientes, es la vía esencial para acumular evidencia científica a favor o en contra de las hipótesis bajo estudio. La replicabilidad, por tanto, está relacionada con la probabilidad de que un estudio independiente produzca resultados consistentes con el estudio original. El concepto de reproducibilidad es una variación reciente de este concepto (Peng, 2015). Se refiere a un estándar de mínimos especialmente referido a la investigación que implica computación masiva de datos. Se define como la posibilidad de recalcular los resultados originales dadas la base de datos y las especificaciones analíticas originales (p.e. código y documentación).

deseado, no informar de ensayos fallidos, eliminación a posteriori de observaciones o variables que no apoyan las hipótesis planteadas. Es decir, un elenco más o menos amplio de prácticas de investigación cuestionables.

Todo lo anterior ha promovido un amplio y vivo debate en el ámbito de la investigación psicológica en los últimos años, del que no daremos cuenta de modo exhaustivo pero que puede ser bien ilustrado a partir de al menos dos indicadores:

- La publicación de numerosas secciones especiales y trabajos de discusión sobre prácticas de investigación cuestionables, replicabilidad y reproducibilidad en reconocidas revistas, tales como *Perspectives on Psychological Science* (2012, 2014) o *Psykometrika* (Sijtsma, 2016; Sijtsma, Veldkamp & Wicherts, 2016; Waldman & Lilienfeld, 2015; Wigboldus & Dotch, 2015).
- El impulso de ambiciosos programas de investigación a gran escala orientados específicamente a replicar resultados de la investigación psicológica publicada, adoptando fundamentalmente nuevas formas de trabajo colaborativo entre equipos de investigación independientes y consorcios de investigación (véase en este sentido Open Science Collaboration, 2015 o Schweinsberg et al., 2016).

En el ámbito de la investigación educativa nos parece que el eco de esta crisis quizá ha sido menor. Ello no quiere decir que posiblemente no se registren problemas similares. Creemos más bien que aparentemente no se ha reunido evidencia o que no se cuenta con tanta información al respecto. Con todo la preocupación no es inexistente. Podemos destacar dos indicadores que nos parecen ilustrativos en este sentido. En primer lugar, la publicación de un trabajo pionero, reciente y muy ambicioso sobre la replicación en las Ciencias de la Educación (Makel & Plucker, 2014). En segundo lugar, el lanzamiento de una nueva revista de la AERA denominada *AERA Open*. En su editorial inaugural la revista se define como abierta a la inspección y la replicación y se hace eco de varios de los elementos asociados al debate de la transparencia y la reproducibilidad sobre los que volveremos más adelante (Warschauer, Duncan & Eccles, 2015).

Un marco general útil para entender el conjunto de preocupaciones que se vienen apuntando es el representado por la meta-investigación,

disciplina o campo de trabajo emergente que se aproxima a la ciencia y sus prácticas “a vista de pájaro”, esto es, adoptando una visión amplia y de conjunto (Ioannidis, Fanelli, Dunne & Goodman, 2015). La meta-investigación se preocupa de cómo se realiza, comunica, verifica, evalúa e incentiva la investigación científica. Y en cada una de esas áreas se focaliza en varios aspectos (ver tabla 1). No abordaremos aquí todos ellos, pero sí varios. Específicamente los vinculados con las prácticas de investigación cuestionables (en el área de métodos), la reproducibilidad y las nuevas prácticas en la publicación y difusión de la investigación científica (en el área de informes). Todo ello con especial atención a las Ciencias Sociales, del Comportamiento y de la Educación. Sobre evaluación y reconocimiento también haremos algunas reflexiones en el apartado de conclusiones<sup>2</sup>.

**TABLA I.** Meta-investigación: áreas temáticas

Area	Intereses específicos ilustrativos
<b>Métodos (realización)</b>	Sesgos y prácticas cuestionables en la realización de investigación y métodos para su reducción; meta-análisis, síntesis de investigación, integración de evidencia; consorcios científicos y equipos colaborativos; integridad investigadora y ética.
<b>Informes (comunicación)</b>	Sesgos y prácticas cuestionables en la publicación y difusión de investigación; gestión de los conflictos de interés; registro de estudios y otros métodos para monitorizar y reducir sesgos y prácticas inadecuadas
<b>Reproducibilidad (verificación)</b>	Obstáculos a las prácticas de datos y métodos compartidos; estudios de replicación; replicabilidad y reproducibilidad de la investigación publicada y métodos para mejorarlas; efectividad de la corrección y auto-corrección de la literatura científica y métodos para mejorarla.
<b>Evaluación (evaluación)</b>	Efectividad, coste y beneficios de nuevas y viejas formas de revisión por pares y otros medios de evaluación de la ciencia y métodos para mejorarlos.
<b>Incentivos (reconocimiento)</b>	Precisión, efectividad, costes y beneficios de viejos y nuevos enfoques para evaluar y ordenar el rendimiento, la calidad y el valor de la investigación de individuos, grupos e instituciones.

Fuente: Adaptado de Ioannidis, Fanelli, Dunne & Goodman (2015)

<sup>2)</sup> Una fuente de información útil sobre meta-investigación en general, con cierto énfasis en el área biosanitaria, es la página del Meta-Research Innovation Center at Stanford – METRICS (<https://metrics.stanford.edu/>). En el ámbito particular de las Ciencias Sociales puede verse: Berkeley Initiative for Transparency in the Social Sciences - BITSS (<http://www.bitss.org/>).

## Sobre sesgos y prácticas de investigación cuestionables

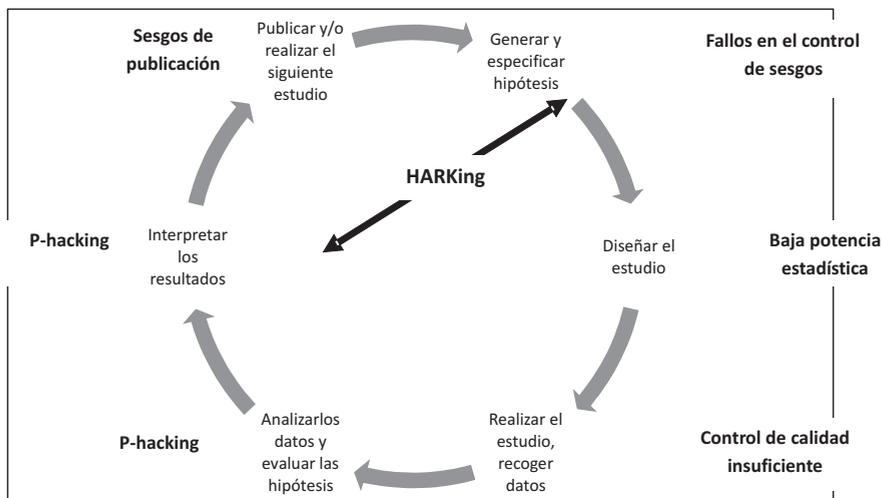
En los últimos años se ha ido generando un cuerpo creciente de literatura sobre sesgos y sobre errores que los investigadores cometen en los procesos de recogida, análisis, archivo y difusión de sus datos y resultados (para una visión de conjunto con énfasis en la investigación psicológica, véase p.e. Sijtsma, 2016 y los artículos de discusión asociados); de modo que en la actualidad se cuenta con una razonable evidencia de la alta prevalencia de algunas prácticas de investigación cuestionables (p.e. para el ámbito de las ciencias cognitivas véase Ioannidis et al., 2014).

En el gráfico I se muestra una versión ideal del modelo hipotético-deductivo del método científico. A cada etapa pueden asociarse amenazas potenciales que, tomadas conjuntamente, pueden contribuir a minar la robustez de la investigación publicada y pueden tener también impacto sobre la capacidad de la ciencia para la autocorrección. Algunos de los fenómenos que han recibido más atención por su alta prevalencia son (Munafò et al., 2017):

- *P-hacking*, también conocido como informe selectivo. Ha centrado gran parte del debate sobre la *crisis estadística de la ciencia*. Ocurre cuando los investigadores intentan varios análisis estadísticos o varias especificaciones sobre la elegibilidad de los datos buscando resultados estadísticamente significativos y, entonces, cuando los hallan en alguno de los múltiples escenarios que se han explorado, sólo se informa de éstos últimos. Las prácticas identificadas como más comunes en este grupo incluyen, entre otras:
  - Prescindir de condiciones experimentales que producen resultados inconsistentes con lo esperado y únicamente analizar e informar de los resultados del subconjunto de condiciones restantes; combinar o separar tratamientos a posteriori, una vez obtenidos los resultados; no informar de todas las variables analizadas, sólo de aquellas asociadas a resultados deseados; reconstruir a posteriori una variable dependiente, por ejemplo usando únicamente algunos ítems, tras probar efectos significativos en esa re-definición operativa y no informar de ello.
  - Eliminar o no casos extremos u otras observaciones en función de las re-evaluaciones de la hipótesis nula; o incluir o excluir covariables a la vista de los resultados y no informar de ello.

- Realizar análisis mientras se recogen los datos de un experimento o estudio y detenerse cuando se obtiene un resultado estadísticamente significativo, esto es, incrementar paso a paso el tamaño de la muestra, evaluando el estadístico de interés en cada paso, hasta que éste alcanza la significación y entonces detener la recogida de datos.
- **HARKing**, término empleado para designar la formulación a posteriori de las hipótesis de investigación: esto es, una vez obtenido el resultado. El escenario más común es informar de un estudio exploratorio como si fuera confirmatorio, en definitiva. Por ejemplo, no informar de las re-especificaciones realizadas sobre el modelo original en el modelo finalmente presentado.
- **Sesgos de publicación**. Los investigadores tienden a no considerar la publicación de trabajos donde no se obtuvieron resultados estadísticamente significativos, lo que está estrechamente vinculado con el ecosistema y la cultura de la publicación científica contemporánea.

GRÁFICO I. Amenazas a la reproducibilidad en el proceso de investigación científica



Fuente: Munafò et al. (2017)

La cuestión, una vez establecida la existencia de tales amenazas para una acumulación fiable y robusta de la evidencia científica, es cómo combatirlas. Las soluciones posibles y las propuestas son múltiples, de naturaleza variada y situadas a distintos niveles en el esquema meta-investigativo que mostramos en la tabla I. En los epígrafes que siguen trataremos de tres aspectos, no exhaustivos pero sí muy incluyentes y muy presentes en la literatura como líneas de acción útiles.

## **Alternativas y “nuevos” énfasis en el análisis estadístico de datos científicos**

En el inventario de prácticas de investigación cuestionables se incluye muy fundamentalmente prácticas que inflacionan el error tipo I, esto es, que aumentan los falsos-positivos, por lo que su análisis está fuertemente asociado al debate de casi siete décadas sobre el contraste de la hipótesis nula (Balluerka, Gómez et al., 2005; Fernández-Cano & Fernández-Guerrero, 2009; Harlow, Mulaik & Steiger, 2016). Aunque creemos que se trata de un tema bien conocido, su mención parece imprescindible, puesto que en el contexto de la crisis de la reproducibilidad el debate ha resurgido con un vigor renovado.

Ha sido muy señalado que los ecosistemas de la producción y la publicación científica gravitan de un modo casi exclusivo sobre el modelo de investigación centrado en el contraste de hipótesis nula y por tanto generan fuertes incentivos para la obtención de resultados estadísticamente significativos y para la publicación exclusiva de los mismos, lo que tiene un vínculo directo con los sesgos de publicación y el informe selectivo. De hecho hay quien ha atribuido directamente la crisis de la reproducibilidad al uso de las pruebas de significación y defienden abierta y sencillamente su abandono. Un representante señalado de esta posición es Cummings, promotor de la llamada *nueva estadística* (Cummings, 2014), que promueve el uso exclusivo de la estimación frente al test de hipótesis y enfatiza tres elementos: el uso de tamaños de efecto, intervalos de confianza y meta-análisis. Las posiciones más extremas han tenido su impacto. El más llamativo, sin duda, lo representa el hecho de que por primera vez una revista de Psicología no marginal, indexada en JCR (*Basic and Applied Social Psychology*) en el año 2015 requiriera expresamente no usar contrastes de hipótesis nula en sus originales (Trafimow & Marks, 2015). También ha resurgido con

intensidad la propuesta de adoptar el enfoque bayesiano de la inferencia estadística (Mulder & Wagenmakers, 2016). Pero estas propuestas, digamos extremas, también han sido ampliamente contestadas desde posiciones más moderadas (y a nuestro juicio particular también más ponderadas) que defienden que el uso correcto de los test de hipótesis y su interpretación cuidadosa son una herramienta clave en el análisis de datos científicos (Morey, Rouder, Verhagen & Wagenmakers, 2014). De hecho se ha calificado de “pista falsa” la exclusiva atención prestada a este asunto en el contexto de la crisis, como si la sola eliminación del contraste de hipótesis nula resolviera todos los problemas y sus alternativas garantizaran por sí solas la validez, robustez y avance efectivo de la ciencia (Savalei & Dunn, 2015).

En este contexto cabe entender la publicación de una reciente declaración institucional de la American Statistical Association sobre el uso de la significación estadística y los p-valores en la investigación científica (Wasserstein & Lazar, 2015). Realmente, como se reconoce en el propio texto, no hay nada nuevo, pero es un buen indicador del estado del debate en el ámbito científico, que parece haber obligado a un pronunciamiento formal de este tipo por parte de la ASA.

En todo caso, como fruto de este renovado debate, se ha dibujado un espacio de amplio consenso sobre la necesidad de revisar algunas prácticas y adoptar definitivamente otras. Entre estas últimas destaca la necesidad de informar siempre del tamaño del efecto y de su precisión, mediante el correspondiente intervalo de confianza, práctica recomendada desde hace más de tres décadas. Y también la de consolidar un razonamiento netamente meta-analítico cuando se aborda la evidencia disponible sobre un problema, interpretando más cabalmente la información limitada que naturalmente puede derivarse de un estudio cualesquiera particular.

El asunto del tamaño del efecto ciertamente no es un tema nuevo. Las recomendaciones en este sentido se remontan a los años 80 (Thompson, 2008), fueron ampliamente difundidas por la APA a fines de los años 90 (Wilkinson & The Task Force on Statistical Inference, 1999), están presentes en todos sus manuales de publicación desde 2001 (p.e. APA, 2001; APA, 2010) y también en los estándares para la publicación de investigación empírica de la AERA (AERA, 2006). Pero realmente se está tratando como si lo fuera. Cabe preguntarse cuál ha sido el grado de adopción de las recomendaciones. Porque quizá ahí esté la explicación de por qué un *tema viejo* sigue siendo actual en algún sentido.

Es interesante en el contexto norteamericano el trabajo de Peng, Chen, Chiang & Chiang (2013). Las autoras compilaron y analizaron 31 revisiones previas que se habían centrado en evaluar el uso e informe del tamaño del efecto en las revistas norteamericanas de Psicología y Educación. A esta minuciosa e ingente revisión de revisiones añaden su propia revisión de 451 artículos publicados en 2009 y 2010. El conjunto de 32 revisiones se dividen en dos periodos: antes y después de 1999 (fecha de publicación del conocido informe de Wilkinson y la TFSI). Se revisaron en total 116 revistas, incluyendo muchas de las asociadas a la APA y las asociadas a AERA.

De un modo apretado puede decirse que este trabajo pone de manifiesto dos cosas. En primer lugar, se identificó globalmente un incremento del uso de los índices del tamaño del efecto, entendido como un aumento en la tasa media de presencia en los artículos de investigación, que sigue siendo no obstante modesta. Aunque se registra una gran variedad entre revistas, las tasas globales pueden ser ilustrativas. Antes de 1999 la tasa media fue igual a 29,9% (mediana 29,4%) y después de 1999 igual a 54,7% (mediana 58%). En segundo lugar, se identificaron unas prácticas aún deficientes en su estimación y en su interpretación, entre las que destaca el escaso uso de los intervalos de confianza para los tamaños del efecto. Parece claro que la adopción de las recomendaciones resulta lenta y no fácil.

En el ámbito de la investigación educativa española no hemos podido encontrar trabajos publicados sobre el uso de medidas de tamaño del efecto y de sus intervalos de confianza. En Psicología sí hemos podido localizar dos trabajos que pueden ofrecer algún indicador siquiera grueso del estado de la cuestión.

García, Campos y De la Fuente (2011) revisaron 787 artículos que se publicaron en cuatro revistas españolas entre los años 2003 y 2008 (*Psicothema*, *Spanish Journal of Psychology*, *Psicológica*, *Internacional Journal of Clinical and Health Psychology*). Se encontró un porcentaje de reporte, respectivamente, del 21%, 32%, 3% y 19% en cada una de las revistas citadas. Como resultado final, y evaluando el total de publicaciones, se tuvo un porcentaje de aplicación del 21%. Caperos y Pardo (2013) analizaron los artículos publicados en 2011 en cuatro revistas españolas de Psicología indexadas en la base de datos JCR (*Anales de Psicología*, *Psicológica*, *Psicothema*, y *Spanish Journal of Psychology*). Sus resultados indican que el 41% de los artículos incluyó alguna medida del tamaño

del efecto. Conviene notar que estas dos revisiones no incluyeron el uso de los intervalos de confianza para el tamaño del efecto, recomendación presente ya en la edición de 2010 de las normas APA y que previsiblemente llevarán un tiempo complementario o adicional.

Parece claro que hay un amplio margen para la mejora. También que no parece plausible encontrar tasas mucho mayores en el ámbito de la investigación educativa.

La prevalencia de sesgos estadísticos, las malas prácticas y la lentitud registradas en la adopción de innovaciones estadísticas y prácticas mejoradas pueden sin duda ser usadas para desacreditar la validez de la ciencia. Sin embargo, creemos que el uso de métodos estadísticos rigurosos y su interpretación cuidadosa constituye una característica esencial, poderosa y sólida de la buena ciencia y una herramienta clave para sustentar su integridad. Y es en este sentido en el que se trata de avanzar.

## **Impulso renovado a la replicación y la reproducibilidad: el caso de las Ciencias de la Educación y del Comportamiento**

En el ámbito de las Ciencias Sociales los trabajos de réplica han sido muy infrecuentes a pesar de su contribución clave al avance del conocimiento científico (Makel & Plucker, 2014). De hecho hasta hace unos pocos años no se contaba con una revisión sistemática de la replicación en investigación educativa.

En el trabajo de Makel & Plucker (2014) ya citado se analizaron las 100 revistas incluidas en la categoría *Education & Educational Research* del JCR con mayor factor de impacto de 5 años (edición 2011) y se analizaron todos los artículos que incluían el término de búsqueda *replicat\** en el texto, considerando la historia completa de cada publicación.

La tasa general de replicación encontrada en el estudio fue igual a 0,13% (221/164.589 artículos), ocho veces menor que la estimada en Psicología por estos mismos autores (Makel, Plucker & Hegarty, 2012). También se identificó una evolución positiva de la misma en perspectiva temporal. Desde 1990 la tasa es casi cuatro veces mayor. O en otras palabras, los estudios de réplica han pasado de ser 1 de cada 2000 artículos de educación a casi 1 de cada 500.

En lo que se refiere al análisis de los resultados, la tasa general de réplicas exitosas fue igual a 67,4%. Estos resultados, que indican que la mayoría de los estudios fueron exitosamente replicados, si bien difieren de los hallados en el ámbito de la Medicina y las Ciencias de la Salud (bastante menos halagüeños), parecen coincidir con los hallados en Psicología. No obstante los autores llaman la atención sobre el hecho de que casi la mitad de las réplicas (48,2%) fueron llevadas a cabo por los mismos equipos que publicaron el estudio original. Y la tasa de éxito en la réplica fue mayor cuando estuvo a cargo del mismo equipo o hubo cierto solapamiento en la autoría que cuando la llevó a cabo un equipo completamente independiente. Conviene entonces considerar las posibles limitaciones asociadas a las auto-réplicas, pues asiste la duda de si la mayor tasa de éxito es fruto del mayor conocimiento y experiencia o si se replican también los sesgos y las prácticas de investigación cuestionables informadas por la literatura en Ciencias Sociales. Esta cuestión requiere por tanto de investigación adicional.

En todo caso parece claro que la replicación es extremadamente infrecuente y que debe ser potenciada. Como señalan los autores de este trabajo, la replicación no es la panacea, no resolverá todos los retos y problemas sobre el rigor, la fiabilidad, la precisión y la validez de la investigación educativa. Sin embargo seguirla ignorando implícita o explícitamente indica una honda falta de comprensión sobre la ciencia y sobre cómo opera su avance.

Estas consideraciones han calado en el ámbito de la Psicología, donde la crisis como señalamos ha sido muy honda y donde se han desarrollado en los últimos años trabajos específicamente centrados en la replicación y la reproducibilidad. Comentaremos brevemente el trabajo ya citado de la *Open Science Collaboration*, publicado en *Science*. Este artículo tuvo y sigue teniendo un enorme impacto (1225 citas en Google Scholar el 1 de septiembre de 2017), ha originado un interesante debate académico con críticas y contra críticas y de algún modo ha venido a constituirse en un clásico en apenas dos años.

Se trata de un ambicioso proyecto colaborativo a gran escala, en la que participaron equipos de investigación independientes de todo el mundo, para obtener una estimación inicial de la reproducibilidad de la ciencia psicológica. Concretamente se llevaron a cabo réplicas de 100 estudios correlacionales y experimentales publicados en 2008 en tres revistas de Psicología de alto impacto (*Psychological Science*, *Social Psychology*,

*Experimental Psychology: learning, memory and cognition*) usando diseños con alta potencia y los materiales originales cuando estuvieron disponibles. Se usaron 5 indicadores para evaluar el éxito de la réplica: significación y valor-p, tamaños del efecto, evaluación subjetiva de los equipos de replicación y meta-análisis de los tamaños del efecto. Los resultados referidos a estos cinco indicadores son complejos y difíciles de resumir en unos párrafos, pues además incluyen análisis de posibles factores asociados a la reproductibilidad y se llevaron a cabo estudios exploratorios para inspirar la investigación futura. Pero algunos datos pueden ser bien ilustrativos. El 97% de los estudios originales obtuvieron resultados significativos frente al 36% de los correspondientes a las réplicas. El tamaño del efecto medio estimado en las réplicas fue igual a .197, la mitad del correspondiente a los estudios originales (.403), lo que señala un descenso sustancial. Finalmente, el 39% de los efectos fueron valorados subjetivamente como replicados por los equipos de investigación.

Tomados en conjunto, estos y el resto de los resultados ofrecen una conclusión clara a juicio de los autores: una gran proporción de réplicas produjeron una evidencia más débil para los resultados originales, a pesar de usar los materiales proporcionados por los autores, de revisar anticipadamente la fidelidad metodológica y de la alta potencia estadística para detectar los tamaños del efecto originales.

En todo caso, como los propios autores apuntan, este trabajo deja muchas cuestiones abiertas, y debe verse como un primer paso, no como un punto de llegada. Los resultados originales ofrecieron evidencia tentativa, las réplicas añaden evidencia adicional, confirmatoria. En algunos casos las réplicas incrementan la confianza en los resultados originales. En otros casos más bien sugieren que se necesita más investigación para establecer la validez de los resultados originales. Y así es como cabalmente deberían ser entendidos estos esfuerzos.

## Ciencia Abierta

El acuerdo unánime en la necesidad de aumentar la transparencia y la apertura en los procesos mediante los que se realiza, difunde y evalúa la investigación científica (como factores clave para garantizar la reproducibilidad) ha movido en los últimos años a diseñar y poner en

marcha iniciativas variadas y que en conjunto vienen conformando lo que se conoce como Ciencia Abierta u *Open Science* (Morey et al., 2016; Nosek et al., 2015). **Una visión de conjunto y una buena ilustración del tipo de iniciativas que se van abriendo paso son las promovidas por el *Center for Open Science*, organización estadounidense sin ánimo de lucro orientada específicamente al desarrollo de herramientas tecnológicas y soportes gratuitos y de libre acceso para la promoción de la ciencia abierta y la reproducibilidad (<https://cos.io/>). Describimos a continuación tres de sus proyectos, especialmente representativos de los avances en este campo.**

### **Directrices para la promoción de la transparencia y la apertura (*Transparency and Openness Promotion- TOP Guidelines*)**

Esta propuesta incluye estándares para ocho módulos que pueden ser adoptados y desarrollados libremente en parte o totalmente por las revistas científicas. Cada módulo además incluye tres niveles de exigencia creciente en términos de transparencia. Se define así un marco flexible que reconoce que no todos los estándares son aplicables a todas las revistas o todas las disciplinas. Los módulos son:

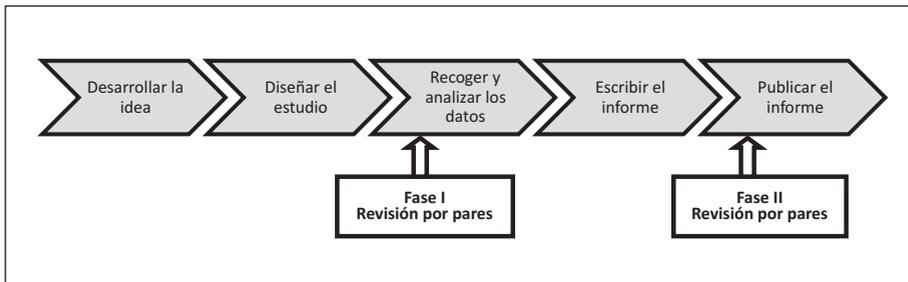
- Estándares de citación
- Transparencia de datos
- Transparencia de métodos analíticos (código)
- Transparencia de materiales de investigación
- Transparencia de análisis y de diseño
- Pre-registro de estudios
- Pre-registro de planes de análisis
- Replicación

Hay un número creciente de revistas adoptando este marco en campos disciplinares variados. Por ejemplo, el 15 de marzo de 2017 se adhirió la revista *Nature* y las revistas asociadas publicadas por Springer y Biomed Central.

Un tema de especial interés nos parece el de los artículos e informes pre-registrados. Constituyen un nuevo formato de publicación que enfatiza la importancia de la pregunta de investigación y la calidad de la

metodología mediante la revisión por pares previa a la recogida de datos (Nosek & Lakens, 2014). Concretamente se trata de aceptar protocolos de investigación de alta calidad provisionalmente, que serán publicados con independencia de los resultados obtenidos si los autores siguen la metodología especificada (ver gráfico II).

GRÁFICO II. Proceso de pre-registro de artículos y planes de análisis



Fuente: Adaptado de Center for Open Science (<https://cos.io/rr/>)

El formato está diseñado para incentivar las buenas prácticas y elimina una variedad de prácticas de investigación cuestionables, tales como la baja potencia estadística, el informe selectivo de resultados o los sesgos de publicación, mientras permite que se complete con flexibilidad el informe con resultados no previstos, calificándolos como tales. Se enfatiza así la distinción entre investigación exploratoria y confirmatoria para una adecuada valoración y evaluación de los resultados de la investigación científica.

En el momento de preparar este artículo más de medio centenar de revistas usan este formato, bien incluyéndolo como una forma regular en su política de aceptación de originales o bien como parte de números especiales.

## Marco de Ciencia Abierta (*Open Science Framework - OSF*)

Entorno *web* abierto y gratuito que permite conectar y dar soporte al flujo de trabajo de la investigación a lo largo de todo el ciclo de un

proyecto, pudiendo ser usado para colaborar, documentar, archivar, compartir y registrar proyectos de investigación, materiales y datos. Hay otras aplicaciones y plataformas, pero esta es una buena ilustración del tipo de herramientas y software innovador desarrollados para dar soporte a las nuevas prácticas de ciencia abierta. Algunas de las revistas que, por ejemplo, recomiendan este marco como repositorio para compartir datos y materiales y para el pre-registro de artículos son: *Biomed Central journals*, *Cognition*, *Nature*, *Psychological Science*, *Perspectives on Psychological Science*, *PloS*, *Science*, *Scientific Data*.

### Distintivos de Ciencia Abierta (*Open Science Badges*)

Se usan para identificar y reconocer expresamente a los artículos que cumplen con determinados estándares de transparencia y apertura. Pueden otorgarse sobre la base de la simple declaración del autor o como resultado de un proceso de revisión por pares, de acuerdo con la política editorial que fije cada revista.

Cuando preparamos este artículo una veintena de revistas habían adoptado estos distintivos. A título de ejemplo, los requisitos que se exigen para otorgar el distintivo de datos abiertos son:

- Proporcionar una URL, DOI, u otro *path* permanente para acceder a un repositorio de acceso público y abierto.
- Proporcionar junto con los datos el conjunto de información suficiente para que un investigador independiente pueda reproducir los resultados informados en el artículo (metadatos).
- Contar con un tipo de licencia abierta que permita a otros copiar, distribuir y hacer uso de los datos aunque igualmente se permita mantener los derechos y el crédito a los propietarios cuando corresponda (*Creative Commons* ha definido recientemente varias licencias con este objetivo).

Con fecha 1 de agosto de 2017 la American Psychological Association – APA anunció un acuerdo con el *Center for Open Science* mediante el cual sus revistas ofrecerán distintivos de ciencia abierta y harán uso del Marco de Ciencia Abierta (OSF) como repositorio para sus datos y para la gestión de servicios de *preprint* (APA, 2017).

## **Conclusiones: algunos retos de futuro para la investigación educativa española**

De la revisión realizada en los epígrafes precedentes cabe derivar algunos retos y desafíos que quizá deba enfrentar la investigación educativa en nuestro país. Vamos a ordenar algunos de ellos sobre cuatro dimensiones.

### **Análisis de las prácticas actuales**

Al preparar este trabajo hemos podido identificar una cierta falta de tradición en la revisión y evaluación de las prácticas de investigación educativa en España, a diferencia de lo que ocurre en otros contextos, o en el nuestro, en otros ámbitos afines. En el ámbito de la Psicología pueden encontrar trabajos que, por ejemplo, analizan el uso de métodos estadísticos para identificar malas y buenas prácticas y elaborar recomendaciones de uso (véase, además de los trabajos ya citados de Caperos & Pardo, 2013 y García, Campos & de la Fuente (2011); el de Izquierdo, Olea y Abad, 2014, entre otros). Por tanto creemos que quizá tenemos aquí una primera cuestión pendiente. Nos falta información o evidencia que permita diagnosticar y en su caso tratar de corregir. El desarrollo de trabajos centrados en analizar las prácticas estadístico-metodológicas actuales en la investigación educativa española constituye entonces una línea de trabajo necesaria y oportuna.

### **Formación, información y sensibilización de los investigadores**

Dada la alta prevalencia con la que se han encontrado prácticas de investigación cuestionables parece razonable preguntarse si realmente en todos los casos los investigadores somos realmente conscientes de incurrir en ellas. Considerar este escenario puede ayudar a entender mejor el asunto. Quizá entre los investigadores se da una infravaloración o falta de conciencia de la dificultad de los temas estadísticos, dificultades para entender, por ejemplo, el impacto sobre las tasas de error tipo I de las prácticas de investigación cuestionables. Centrarnos en la intencionalidad para distinguir la mala praxis (por falta de información o formación) del

fraude, como ha sido señalado en la literatura (Sijtsma, 2016), puede sugerir la necesidad de intervenir sobre los procesos de formación inicial y permanente de los investigadores en el ámbito estadístico-metodológico. Parece entonces pertinente una cierta reflexión sobre la formación de postgrado inicial y también sobre la oportunidad de arbitrar cauces de formación permanente para los investigadores en activo. Porque las dificultades y la lentitud en la adopción de innovaciones y buenas prácticas renovadas por parte de los investigadores es un asunto que ha recibido una atención creciente (Henson, Hull & Williams, 2010; Sharpe, 2013; Cohen, 2017) y creemos que quizá requiera de consideración también en nuestro contexto.

### **Políticas editoriales: el papel de editores y revisores**

En todo caso los investigadores queremos publicar y por tanto puede pensarse que estamos bien dispuestos a responder a los requisitos que ello exige. De ahí que en principio las políticas editoriales de las revistas científicas tengan un papel esencial en el modo en que se realiza y se difunde la investigación. Es razonable pensar que si las revistas hacen explícitos unos adecuados estándares estadístico-metodológicos para la aceptación de originales, asumidos y tomados como referente común por los revisores, no penalizan o invitan a la publicación de réplicas y fomentan o instauran políticas de ciencia abierta, ello tendrá un efecto positivo directo, como parece estar poniéndose de manifiesto en otros contextos (véase Munafò et al., 2017). En la práctica creemos que ello supone un reto fabuloso en nuestro entorno y en nuestro ámbito. En esta década, en nuestro país, hemos asistido a una consolidación extraordinaria de un buen número de revistas científicas de calidad en educación, con grandes y frecuentemente desinteresados esfuerzos por parte de la comunidad investigadora y muy especialmente de editores y revisores (Ruiz-Corbella, Galán & Diestro, 2014). Un buen número de publicaciones han ingresado en índices internacionales y sólo mantener los logros habidos creemos personalmente que ya es en sí un gran reto. No obstante, creemos que las líneas apuntadas señalan un camino de futuro posiblemente necesario que ya han emprendido importantes revistas de educación (véase p.e. Lopez, Valenzuela, Nussbaum & Tsai, 2015).

## **Políticas académicas y científicas: el papel de los financiadores y evaluadores de la producción investigadora**

Por último, pero sin duda no menos importante, creemos que es esencial considerar el papel clave que tienen los agentes que financian y evalúan la producción científica de investigadores, equipos e instituciones, porque a la postre son los que diseñan el sistema de incentivos y reconocimiento que modula en gran medida las conductas individuales y colectivas. Si se prima en la evaluación de la producción exclusivamente la cantidad y la presión por publicar sigue caracterizando extraordinariamente el ecosistema en el que se desarrolla el trabajo de los investigadores, parece difícil que se puedan mejorar algunas de las prácticas citadas. Si publicar datos en abierto no es incentivado de ningún modo, posiblemente no será una práctica que se extienda fácilmente; o si la participación en proyectos de amplio alcance colaborativos y los consorcios de investigación no se incentivan, resultará difícil que se introduzcan y consoliden en la práctica. Coincidimos por tanto con aquéllos que han señalado que los cambios en este nivel institucional son clave (Smaldino & McElreath, 2016) y es donde de hecho se concentran gran parte de los retos y desafíos que se han venido apuntando a lo largo de este artículo.

### **Reflexión final**

En esencia lo que hemos querido subrayar en este trabajo es que el progreso científico es un proceso acumulativo de reducción de la incertidumbre (más que de la obtención de certezas) y que este proceso sólo puede tener éxito si la ciencia misma opera sobre la base de un natural y sistemático escepticismo sobre sus propios hallazgos. Creemos que el análisis hecho pone de manifiesto que hay un claro margen para la mejora de las prácticas científicas. Pero también esperamos haber puesto suficientemente de manifiesto que ante las hipótesis, más que plausibles, de que el ecosistema de la ciencia y su cultura actual pueden estar afectando negativamente a la validez y reproducibilidad de sus resultados, el sistema científico se está comportando como cabía esperar: revisando sistemática y críticamente sus prácticas, formulando alternativas de mejora y diseñando nuevos mecanismos para la auto-corrección. Una

mirada como la que hemos propuesto aquí creemos que puede ser útil para contribuir a la reflexión sobre nuestras propias prácticas y al avance, en definitiva, de la investigación educativa en nuestro país.

## Referencias bibliográficas

- American Educational Research Association. (2006). Standards for reporting on empirical social science research in AERA publications. *Educational Researcher*, 35(6), 33–40. doi:10.3102/0013189X035006033.
- American Psychological Association. (2001). *Publication manual of the American Psychological Association* (5th ed.). Washington, DC: American Psychological Association.
- American Psychological Association. (2010). *Publication manual of the American Psychological Association* (6th ed.). Washington, DC: American Psychological Association.
- American Psychological Association. (2017, August 1). *APA Journals Program collaborates with Center for Open Science to advance open science practices in psychological research*. Retrieved from: <http://www.apa.org/news/press/releases/2017/08/open-science.aspx>.
- Baker, M. (2016). 1,500 scientists lift the lid on reproducibility. *Nature*, 533(7604), 452–454. doi:10.1038/533452a
- Balluerka, N., J. Gómez, et al. (2005). The controversy over null hypothesis significance testing revisited. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences* 1(2): 55–70.
- Caperos, J.M., & Pardo, A. (2013). Consistency errors in p-values reported in Spanish psychology journals. *Psicothema*, 25(3), 408–414. doi: 10.7334/psicothema2012.207.
- Cohen, B. H. (2017). Why the resistance to statistical innovations? A comment on Sharpe (2013). *Psychological Methods*, 22(1), 204–210. doi: 10.1037/met0000058.
- Cumming, G. (2014). The new statistics: why and how. *Psychological Science*, 25, 7–29. doi:10.1177/0956797613504966.
- Fernández-Cano, A., & Fernández-Guerrero, I. (2009). *Crítica y alternativas a la significación estadística en el contraste de hipótesis*. Madrid: La Muralla.

- García, J. Campos, E., & De la Fuente, L. (2011). The use of the effect size in JCR Spanish journals of Psychology: from theory to fact. *The Spanish Journal of Psychology*, *14*(2), 1050-1055.
- Goodman, S.N. (2016). Aligning statistical and scientific reasoning. Misunderstanding and misuse of statistical significance impede science. *Science*, *352* (6290), 1180-1181. doi: 10.1126/science.aaf5406.
- Harlow, L. L., Mulaik, S. A., & Steiger, J. H. (Eds.). (2016). *What if there were no significance tests?* Classic Edition. New York: Routledge.
- Henson, R.K., Hull, D.M., & Williams, C.S. (2010). Methodology in our education research culture: toward a stronger collective quantitative proficiency. *Educational Researcher*, *39*(3), 229-240. doi: 10.3102/0013189X10365102.
- Ioannidis, J. P., Fanelli, D., Dunne, D. D., & Goodman, S. N. (2015). Meta-research: evaluation and improvement of research methods and practices. *PLoS Biology*, *13*(10),1-7. doi: 10.1371/journal.pbio.1002264.
- Ioannidis, J. P., Munafo, M. R., Fusar-Poli, P., Nosek, B. A., & David, S. P. (2014). Publication and other reporting biases in cognitive sciences: detection, prevalence, and prevention. *Trends in cognitive sciences*, *18*(5), 235-241.
- Izquierdo, I., Olea, J., and Abad, F.J. (2014). Exploratory Factor Analysis in validation studies: uses and recommendations. *Psicothema*, *26*(3), 395-400.
- Ledgerwood, A. (2014). Introduction to the Special Section on Advancing Our Methods and Practices *Perspectives on Psychological Science*, *9*(3) 275–277. doi: 10.1177/1745691614529448.
- Lopez,X., Valenzuela, J., Nussbaum, M. & Tsai, C.(2015). Some recommendations for the reporting of quantitative studies (Editorial). *Computers & Education*, *91*,106-110.
- Makel, M. C., & Plucker, J. A. (2014). Facts are more important than novelty: Replication in the education sciences. *Educational Researcher*, *43*, 304 –316. doi: 10.3102/0013189X14545513.
- Makel, M. C., Plucker, J. A., & Hegarty, B. (2012). Replications in psychology research: How often do they really occur? *Perspectives in Psychological Science*, *7*, 537–542. doi:10.1177/1745691612460688.
- McNutt, M. (2014). Journals unite for reproducibility (Editorial). *Science*, *346*(6210), 679-679. doi:10.1038/515007a.
- Morey, R. D., Rouder, J. N., Verhagen, J., & Wagenmakers, E. J. (2014). Why hypothesis tests are essential for psychological science: a comment on Cumming (2014). *Psychological science*, *25*(6), 1289-1290.

- Morey, R.D. et al. (2016). The Peer Reviewers' Openness Initiative: incentivizing open research practices through peer review. *Royal Society Open Science*, 3, 150547. doi: <http://dx.doi.org/10.1098/rsos.150547>.
- Mulder, J. & Wagenmakers, E.J. (2016). Editors' introduction to the special issue "Bayes factors for testing hypotheses in psychological research: Practical relevance and new developments". *Journal of Mathematical Psychology*, 72, 1-5.
- Munafò, M. R., Nosek, B. A., Bishop, D. V., Button, K. S., Chambers, C. D., du Sert, N. P., ... & Ioannidis, J. P. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, 1, 0021. doi:10.1038/s41562-016-0021.
- Nature (2014). Journals unite for reproducibility (Editorial). *Nature*, 515(7525), 7.
- Nosek, B. A., et al. (2015). Promoting an open research culture. *Science*, 348, 1422-1425. DOI: 10.1126/science.aab2374.
- Nosek, B. A. & Lakens, D. (2014). Registered reports. A method to increase the credibility of published results. *Social Psychology*, 45, 137-141. doi: 10.1027/1864-9335/a000192.
- Open Science Collaboration (2015). Estimating the reproducibility of psychological science. *Science*, 349 (6251). doi: 10.1126/science.aac4716.
- Peng, C-Y., Chen, L-T., Chiang, H-M., & Chiang, Y-C. (2013). The impact of APA and AERA guidelines on effect size reporting. *Educational Psychology Review*, 25, 157-209. doi: 10.1007/s10648-013-9218-2.
- Peng, R.D. (2015). The reproducibility crisis in science. A statistical counterattack. *Significance*, 30-32
- Perspectives on Psychological Science. (2012). Special section on replicability in psychological science: A crisis of confidence? Retrieved from: <http://pps.sagepub.com/content/7/6.toc>
- Perspectives on Psychological Science. (2014). Special section on Advancing our methods and practices. Retrieved from: <http://journals.sagepub.com/toc/ppsa/9/3>.
- Ruiz-Corbella, M., Galán, A. & Diestro, A. (2014). Las revistas científicas de Educación en España: evolución y perspectivas de futuro. *RELIEVE*, 20 (2), art. M1. doi: 10.7203/relieve.20.2.436.
- Savalei V., Dunn E. (2015). Is the call to abandon p-values the red herring of the replicability crisis? *Frontiers in Psychology*, 6, 245. doi: 10.3389/fpsyg.2015.00245.

- Schweinsberg, M. et al. (2016). The pipeline project: Pre-publication independent replications of a single laboratory's research pipeline. *Journal of Experimental Social Psychology*, 66, 55-67. doi: <https://doi.org/10.1016/j.jesp.2015.10.001>.
- Sharpe, D. (2013). Why the resistance to statistical innovations? Bridging the communication gap. *Psychological Methods*, 18(4), 572. doi: 10.1037/a0034177.
- Sijtsma, K. (2016). Playing with data—Or how to discourage questionable research practices and stimulate researchers to do things right. *Psychometrika*, 81(1), 1-15. doi:10.1007/s11336-015-9446-0.
- Sijtsma, K., Veldkamp, C.L.S. & Wicherts, J.M. (2016). Improving the conduct and reporting of statistical analysis in Psychology. *Psychometrika*, 81(1), 33-38. doi:10.1007/s11336-015-9444-2.
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, 22, 1359–1366. doi:10.1177/0956797611417632.
- Smaldino, P.E., McElreath, R. (2016). The natural selection of bad science. *Royal Society Open Science*, 3, 160384. doi:<http://dx.doi.org/10.1098/rsos.160384>.
- Thompson, B. (2008). Computing and interpreting effect sizes, confidence intervals, and confidence intervals for effect sizes (pp. 246-262). En Osborne, J.W. (2008)(Ed.). *Best Practice for Quantitative Methods*. London: Sage.
- Trafimow, D. & Marks (2015). Editorial. *Basic and Applied Social Psychology*, 37 (1), 1–2.
- Waldman, I. D., & Lilienfeld, S. O. (2015). Thinking about data, research methods, and statistical analyses: Commentary on Sijtsma's (2014) "Playing with data". *Psychometrika*, 81(1), 16-26. doi:10.1007/s11336-015-9447-z.
- Warschauer, M., Duncan, G. J., & Eccles, J. S. (2015). Inaugural Editorial: what we mean by "open". *AERA Open*, 1 (1),1-2. doi: 10.1177/2332858415574841.
- Wasserstein, R. L., & Lazar, N. A. (2016). The ASA's statement on p-values: context, process, and purpose. *American Statistician*, 70(2), 129-133. doi: 10.1080/00031305.2016.1154108.
- Wigboldus, D. H. J., & Dotch, R. (2015). Encourage playing with data and discourage questionable reporting practices. *Psychometrika*, 81(1), 27-32. doi:10.1007/s11336-015-9445-1.

Wilkinson, L., & The Task Force on Statistical Inference. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist*, 54, 594–604. doi:10.1037/0003-066X.54.8.594.

**Información de contacto:** Ángeles Blanco Blanco. Universidad Complutense de Madrid, Facultad de Educación, Departamento de Investigación y Psicología en Educación. Rector Royo Villanova s/n 28040 Madrid. E-mail: ablancob@ucm.es



# Capital Social y Redes Sociales de Maestros: Revisión Sistemática<sup>1</sup>

## Social Capital and Social Networks of Teachers: Systematic Review

DOI: 10.4438/1988-592X-RE-2017-381-387

Susana López Solé  
Mireia Civís Zaragoza  
Jordi Díaz-Gibson  
*Universidad Ramón Llull*

### Resumen

Estudios recientes demuestran la importancia y la aplicabilidad del concepto del capital social en educación, con especial interés en los maestros, centrándose en el análisis de las estructuras sociales para comprender cómo las relaciones entre ellos pueden facilitar o impedir el acceso a diferentes recursos. Con este propósito, el presente estudio realiza una revisión sistemática de las investigaciones que se han realizado a nivel internacional en relación al capital social y a las redes sociales de maestros para conocer y demostrar la importancia de estos constructos en la mejora de la praxis educativa, así como para fomentar su investigación. La búsqueda en las bases de datos Eric y Web of Science durante el periodo 2000-2015, reporta 143 estudios que cumplen con los criterios de inclusión definidos y, dentro de éstos, 14 trabajos entre los años 2008-2015 que pertenecen a la figura de los maestros, en los que se profundiza en este estudio. La mayoría de los estudios de maestros exploran la mejora del apoyo y del desarrollo profesional, a través de las interacciones y la creación de comunidades profesionales de aprendizaje. En los resultados

---

<sup>1</sup> Con el soporte de Fondos Sociales Europeos y de la *Secretaria d'Universitats i Recerca* del Departamento de *Economia i Coneixement de la Generalitat de Catalunya* (2015 FI\_B 00948, 2016 FI\_B1 00162, 2017 FI\_B2 00142).

se propone un modelo de clasificación de los 14 estudios en relación al capital social. Concretamente, se clasifica según su conceptualización, en base a sus diferentes dimensiones estructurales, relacionales y cognitivas, y su medición, teniendo en cuenta los recursos y las posiciones en la red. Sin embargo, los resultados apuntan a la necesidad de mayor consenso, debido a la heterogeneidad y variabilidad de los estudios revisados. Finalmente, en los estudios existe un acuerdo para explorar el capital social de maestros a partir del análisis de redes sociales con el fin de analizar tanto las redes formales, a partir del soporte intencionado proporcionado por las escuelas, como informales, que aparecen de manera espontánea de las relaciones personales.

*Palabras clave:* capital social, redes sociales, educación, maestros, revisión sistemática

### **Abstract**

Recent studies demonstrate the importance and applicability of the concept of social capital in education, especially in relation to teachers, focusing on the analysis of social structures to understand how relationships between them can facilitate or inhibit access to different resources. With this purpose in mind, the present study performs a systematic review of the research that has been carried out on an international level in relation to teachers' social capital and social networks to understand and demonstrate the importance of these constructs in the improvement of the educational praxis, as well as to promote their research. A search in the Eric and Web of Science databases during the period 2000-2015 reports 143 studies that meet the defined inclusion criteria and, of these, 14 studies between the years 2008-2015 that focus on the figure of the teacher, which this study analyses in greater depth. Most teacher-focused studies explore the improvement of teacher support and professional development through interactions and the creation of professional learning communities. The results propose a classification model of the 14 studies in relation to social capital. Specifically, it is classified according to its conceptualization, based on its different structural, relational and cognitive dimensions, and its measurement, taking into account resources and positions in the network. However, results point to the need for greater consensus, due to the heterogeneity and variability of the studies reviewed. Finally, in the studies there is an agreement to explore the social capital of teachers from the analysis of social networks in order to analyse both formal networks, from the intentional support provided by schools, and informal networks, which appear spontaneously from personal relationships.

*Key words:* social capital, social networks, education, teachers, systematic review

## Introducción

Actualmente formamos parte de una sociedad compleja caracterizada por los continuos y acelerados cambios que constituyen una realidad en nuestras vidas (Bauman, 1999; Morin, 2004). En las últimas décadas, hemos experimentado fuertes cambios sociales en el ámbito económico, laboral o familiar que han provocado una serie de transformaciones en el proceso de socialización de las personas (Bolívar, 2006; Por ello, las interconexiones tienen un papel fundamental para adaptarnos a los cambios que llevarían a la fragmentación y al aislamiento, entendiendo el mundo de manera sistémica.

La preocupación por la innovación pedagógica que vivimos a día de hoy responde en gran medida a la búsqueda para dar respuesta a los cambios que experimenta nuestra sociedad. En este sentido, cada vez más se está optando por el lado *social* de la educación para afrontar los retos educativos actuales, dado el potencial que tienen los vínculos y las interacciones, siendo el capital social su máxima expresión.

Ante la creciente diversidad que hay en las aulas y en el entorno (Clayton, 2007) el maestro tiene un papel fundamental para obtener el éxito escolar y las mejoras educativas en su conjunto (Piqué, Comas y Lorenzo, 2010; Vaillant y Manso, 2013). Así pues, garantizar que los maestros desarrollen su labor en este contexto social y educativo pasa a ser una prioridad estratégica para la educación y es por ello que cada vez más se incrementa la complejidad de su misión y los requerimientos sobre sus competencias y profesionalidad (Darling-Hammond, 2006; Schleicher, 2012).

## Teoría y concepto del capital social

Adoptando la definición de Lin (1999) el capital social hace referencia a los recursos incrustados en una estructura social que son accesibles o movilizados por una acción intencional. En las últimas décadas, el capital social ha emergido como uno de los conceptos destacados en el ámbito de las ciencias sociales (Lin, 1999; Daly, 2010; Dika y Singh, 2012).

Encontramos su origen a finales de los años 1980 con las teorías de los sociólogos Bourdieu (1986) y Coleman (1988). Posteriormente, un gran número de teóricos han escrito sobre el concepto del capital social en sus diversas formas y contextos como Granovetter (1973), Putnam

(1995), Portes (1998), Lin (1999), Fukuyama (2001) y Burt (2004). De sus distintas aproximaciones, se desprende que los principales componentes que conforman el capital social se pueden sintetizar en: redes, confianza, normas y sanciones (Baron, Field y Schuller, 2000; Halpern, 2005). Las redes y la confianza suelen ser los dos elementos clave para definir el concepto (Baron, Field y Schuller, 2000). La red hace referencia a los vínculos entre los diferentes nodos interconectados que conforman la estructura social (Castells, 1996; Scott, 2000). Por otro lado, la confianza ha sido estudiada por diferentes autores (Coleman, 1988; Putnam, 1995; Fukuyama, 1995) como facilitador de la cooperación y del orden social que puede inducir a los actores de la red a la reciprocidad. Asimismo, las normas y sanciones suelen ser elementos frecuentes para definir el capital social y atribuidos a los valores, expectativas y recompensas que se desprenden de las relaciones (Halpern, 2005).

A su vez, se pueden distinguir tres funciones del capital social según las características de sus vínculos: *bonding*, *bridging* y *linking* (Putnam, 2000; Woolcock, 2001; Halpern, 2005). La función *bonding* se refiere a aquellas relaciones más intensas y cercanas entre las personas, como podría ser el caso de los familiares o amigos más cercanos. Así, el *bonding* refuerza las identidades de las personas y la formación de grupos más homogéneos (Halpern, 2005) y se relaciona con el concepto de lazo fuerte o *strong ties* que juega un papel clave en el soporte y bienestar emocional de las personas (Granovetter, 1973). Por otro lado, el *bridging* hace referencia a aquellas relaciones que son más distantes y no tan cercanas, como el caso de amigos de amigos. Este tipo de capital social también se caracteriza por las relaciones entre personas que tienen diferentes orígenes y experiencias (Miller, 2012). Por lo tanto, este tipo de relaciones permiten el acceso a nuevos recursos y oportunidades que no podrían estar disponibles en el caso de las relaciones *bonding*. El capital social *bridging* se puede vincular con los lazos débiles o *weak ties* que propone Granovetter (1973), ya que estas relaciones resultan de mucha utilidad para las personas con tal de poder acceder a informaciones, oportunidades o trabajos. Por último, el capital social de tipo *linking* describe las relaciones verticales entre las personas e instituciones que se encuentren en posiciones que no son similares y que tienen una distribución desigual de recursos y de poder (Halpern, 2005), como sería el caso de acceder a servicios gracias a las conexiones con estructuras e instituciones que tengan mayor poder (Hawkins y Maurer, 2010).

Finalmente, existe una concepción del capital social a través de diferentes dimensiones: estructural, relacional y cognitiva (Nahapiet y Ghoshal, 1998). La dimensión estructural hace referencia a las redes y a los patrones de las conexiones entre los diferentes actores de la red (Nahapiet y Ghoshal, 1998). Por lo tanto, aquellos individuos que se encuentren en posiciones más céntricas en las redes sociales tendrán más acceso a los recursos, junto con las redes que sean más densas en las que los individuos se encuentren más conectados. Por otro lado, la parte relacional se vincula con la confianza y con la calidad de los vínculos, a partir de las normas, valores y expectativas que comparten los individuos de la red (Moolenaar, 2010; Tschannen-Moran, 2014). En esta dimensión, el papel de la confianza juega un papel esencial, puesto que se trata de un elemento relacional que existe en las relaciones sociales con tal de mantener las obligaciones y expectativas mutuas (Liou y Daly, 2014). Finalmente, la parte cognitiva se basa en el conocimiento que se comparte y se intercambia gracias a la comunicación entre los diferentes miembros que componen la red y que comparten un lenguaje y unas narrativas en la red (Liou y Daly, 2014). Así pues, el capital social cognitivo los conforman los recursos que proveen unas representaciones e interpretaciones compartidas y un sistema de significados entre los miembros de la red, como normas y valores (Nahapiet y Ghoshal, 1998).

En este sentido, un mayor capital social de maestros significará un mayor desempeño de su actividad docente (Dika y Singh, 2002; Smethem, 2007; Liou et al., 2016) por lo que investigaciones siguieren la importancia de las redes de colaboración de maestros para incrementar sus competencias y el éxito de sus estudiantes (Penuel, Riel, Krause y Frank, 2009; Daly y Finnigan, 2010).

A finales de los años 90 se produjo un auge del uso del “capital social” en los artículos científicos en un rango amplio de disciplinas (Halpern, 2005). En la actualidad es un tema cada vez más estudiado, también en el ámbito educativo, pero al mismo tiempo más difuso, dada la variación que existe según la aproximación que adopten los distintos autores.

## **Teoría y análisis de las redes sociales**

De la definición de capital social se desprende que puede ser medido y capturado a través de los recursos presentes en las redes sociales, donde

los diferentes actores pueden acceder y apropiarse de estos recursos. Una red consiste en un conjunto de actores (personas, organizaciones, equipos...) que se encuentran conectados por un determinado tipo de vínculo (Kilduff y Tsai, 2003; Daly, 2010).

En este sentido, el interés hacia la teoría y el análisis de las redes sociales se ha incrementado notablemente en los últimos 20 años (Moolenaar, 2010; Carolan, 2013). En el ámbito educativo, este análisis ha permitido estudiar una gran variedad de fenómenos, como las reformas políticas (Coburn y Russell, 2008; Penuel, et al., 2009), el liderazgo educativo (Moolenaar, Daly y Slegers, 2010; Daly y Finnigan, 2011) o las comunidades profesionales de aprendizaje (Slegers, den Brok, Verbiest, Moolenaar y Daly, 2013; Rienties y Kinchin, 2014).

El análisis de redes sociales de maestros permite explorar diferentes aspectos en relación a su formación inicial (Liou et al., 2016; Civiş, López y Díaz-Gibson, 2017) y la colaboración entre profesionales para mejorar su desarrollo profesional (Baumard y Starbuck, 2005; Penuel, Sun, Frank y Gallagher, 2012), entre otros, centrándose en comprender cómo las relaciones entre los actores de una red pueden facilitar o impedir el acceso a los recursos (Cross, Borgatti, y Parker, 2002; Borgatti y Foster, 2003).

Los recursos que se pueden obtener pueden ser de diferente naturaleza, según si se trata de redes instrumentales o expresivas (Lin, 1982; Ibarra, 1993). Las redes instrumentales hacen referencia a aquellas redes que están relacionadas con el trabajo y el desarrollo académico y que implica el intercambio de información, de conocimiento, de consejo, etc. Por otro lado, las redes expresivas están orientadas a la transferencia de recursos afectivos y emocionales como por ejemplo la amistad. Estas redes expresivas contienen un nivel más elevado de confianza, comparada con las instrumentales, y tienden a ser más estables en el tiempo (Ibarra, 1993; Uzzi, 1997).

Asimismo, el análisis de redes sociales se puede clasificar según los diferentes niveles de estudio: el nivel egocéntrico o personal de cada actor, y un nivel sociocéntrico o de estructura completa de toda la red (Borgatti, Everett y Johnson, 2013; Crossley et al., 2015; Hâncean, Molina y Lubbers, 2016). Del mismo modo, existen tres niveles de análisis para realizar un análisis de redes sociales, según si se trata de un análisis sobre el nodo (un actor), la díada (dos actores) o la red (todos los actores de la red) (Daly, 2010; Borgatti, Everett y Johnson, 2013).

Respecto a sus medidas de análisis, a nivel general de la red encontramos las medidas estructurales como la densidad, el tamaño de la red y la centralización global que nos aportan información sobre la red respecto a su cohesión y conectividad (Carolan, 2013; Borgatti, Everett y Johnson, 2013). Pasando al nivel de actor, hay diversas medidas que nos aportan información sobre la posición de un determinado actor respecto a los demás actores de la red. Las medidas de centralidad (Freeman, 1979; Wasserman y Faust, 1998) son las más utilizadas en este nivel de análisis, dentro de las cuales las más frecuentes son: centralidad de grado, centralidad de intermediación y la centralidad de cercanía.

La centralidad de grado hace referencia al número de vínculos de un actor con los demás actores, distinguiendo entre los vínculos que envía un actor hacia los demás (grado de salida) y los vínculos que un actor recibe de los demás (grado de entrada). Por su parte, la centralidad de intermediación mide el grado en el que los actores controlan o median la relación entre pares de actores que no están conectados entre ellos. Esta medida de centralidad está relacionada con la posición de *broker* o puente que juega un papel relevante en la estructura de la red, especialmente en los *structural holes* como resultado de lazos débiles o ausentes en las conexiones entre individuos (Burt, 2004). Por último, la centralidad de cercanía mide la distancia media de las distancias más cortas desde un actor hacia los otros actores.

## Método

### Objetivos

El presente estudio tiene como objetivo general realizar una revisión sistemática del capital social y de las redes sociales de maestros a partir de las investigaciones a nivel internacional sobre esta temática entre los años 2000 y 2015. Asimismo, el estudio persigue los siguientes objetivos específicos:

- Conocer y demostrar la importancia de los constructos del capital social y de las redes sociales en la mejora de la praxis educativa, entrándose en la figura de maestros

- Aproximarse a la teoría (y análisis) de redes sociales para comprender la relevancia del capital social de maestros
- Realizar una propuesta de conceptualización y medición del capital social de maestros

## Estrategia de la búsqueda

Se ha realizado una búsqueda exhaustiva de la literatura en la base de datos ERIC de referencia en el ámbito educativo (Marquès, 1999; Carolan, 2013) y en la colección principal de la base de datos Web of Science de referencia en el ámbito científico (McVeigh y Bates, 2010). Se ha seleccionado ERIC por proporcionar el acceso a un gran número de registros bibliográficos, de artículos de revista y de otros materiales en relación con la investigación educativa. Por otro lado, Web of Science contiene una gran colección de bases de datos bibliográficos, citas y referencias de publicaciones científicas de cualquier disciplina de conocimiento, ofreciendo una mirada transversal y multidisciplinar.

El periodo de recogida de artículos fue de 01.07.15 al 12.11.15 y para la búsqueda se seleccionaron los artículos publicados entre los años 2000 y 2015 utilizando los descriptores: “*network analysis*”, “*social networks*”, “*social capital*”, “*educat\**”, “*school\**” “*and teach\**”. Se seleccionaron los años 2000 y 2015 por la tendencia creciente en estos años de las publicaciones dentro del ámbito educativo en relación a las redes sociales y al capital social (Carolan, 2013; Moolenaar, 2010).

En las bases ERIC y Web Of Science se ha realizado una búsqueda avanzada donde se combinan más de dos descriptores (“DE” en ERIC) o temas (“TS” en Web Of Science). En ambas bases se utilizan los operadores booleanos para buscar registros que incluyan todos los términos separados (AND) o bien para buscar registros que incluyan cualquiera de los términos separados (OR).

El resultado de la búsqueda fue el siguiente: (DE “Network Analysis” OR DE “social networks”) AND DE “Social capital” AND (educat\* OR school\* OR teach\*). Mientras que en la base de datos Web Of Science realizamos la misma búsqueda adaptada a la base de datos (“TS” en vez de “DE”).

Los términos “Network Analysis” y “Social networks” nos permiten aproximarnos a la literatura basada en las redes y en las interacciones.

Por otra parte, el término “Social capital” de acuerdo con la literatura, se refiere a los recursos obtenidos a partir de las redes. Por último, los términos “educat”, “school”, y “teach” nos permiten contextualizar en el ámbito educativo. De manera más específica, estos últimos términos nos permiten acceder a diversos estudios dentro del terreno educativo (“educat”) que profundicen en el ámbito escolar u organizativo (“school”), así como en la figura de los maestros y los procesos de enseñanza (“teach”).

## Criterios de inclusión

Los criterios de inclusión que se establecieron para la selección de los artículos, en base a Liberati et al., (2009), son los siguientes:

- Estudios empíricos con datos recolectados mediante métodos cuantitativos, cualitativos o mixtos.
- Estudios contextualizados en el ámbito educativo y con el fin de mejorar la práctica educativa.
- Estudios que analizan y exploran el concepto del capital social en base a la teoría (y análisis) de redes sociales.
- Estudios que estén publicados con un sistema de revisión por pares (*peer review*) para asegurar el estándar de calidad y fiabilidad.
- Estudios publicados entre los años 2000 y 2015.

## Procedimiento

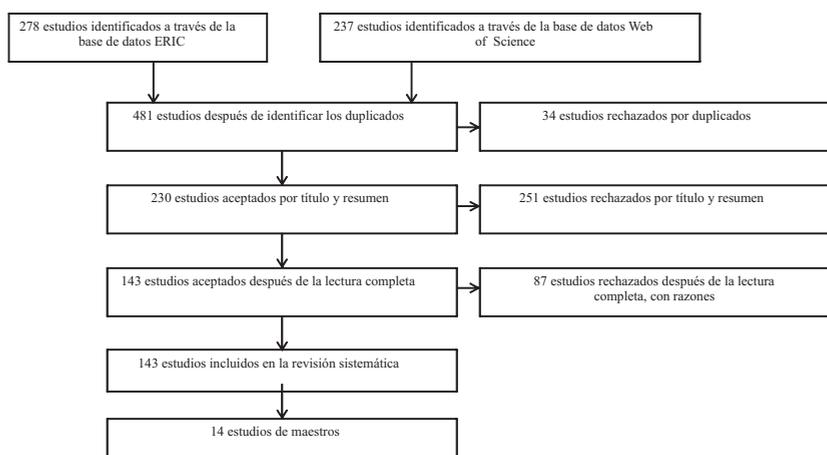
Con el fin de ilustrar el proceso, en el gráfico I se detalla el procedimiento seguido y las distintas fases para realizar la revisión sistemática:

Una vez se identificaron los estudios duplicados de las bases de datos (34), se procedió a la revisión de los 481 estudios para valorar su inclusión, en base a la lectura de títulos y resúmenes. La valoración de los estudios fue dicotómica y cualitativa, con calificaciones de “aceptado” o “rechazado”, de acuerdo con la presencia o ausencia de los criterios definidos. En esta fase se rechazaron 251 estudios por no cumplirlos.

A continuación, se procedió a realizar la lectura completa de los 230 estudios restantes, para aceptar o rechazar mediante un análisis exhaustivo

del conjunto del texto. Concretamente, durante el proceso se rechazaron 87 estudios debido principalmente a que no se contextualizaban propiamente en el ámbito educativo ni su foco principal se centraba en el capital social para la mejora de la praxis educativa. Finalmente se obtuvieron 143 estudios que sí cumplían con los criterios de inclusión para la revisión sistemática, dentro de los cuales 14 pertenecían a maestros.

GRÁFICO I. Distintas fases de la revisión sistemática



Fuente: Adaptación de PRISMA (Liberati et al., 2009)

Toda discordancia durante el procedimiento se resolvió por consenso entre cinco investigadores. Además, de cada uno de los trabajos se extraía la siguiente información de manera detallada para su posterior análisis y discusión:

- Autores y año de publicación
- Contexto geográfico del estudio
- Muestra del estudio (número de participantes y descripción)
- Temática principal del estudio
- Aproximación a la teoría del capital social y a la teoría (y análisis) de redes sociales
- Principales resultados y conclusiones
- Comentarios y observaciones

## Resultados

Los resultados obtenidos a partir de los 14 estudios de maestros apuntan principalmente hacia el capital social como medida de su apoyo e incremento de su desarrollo profesional.

En cuanto a su aproximación a la teoría de redes sociales, los estudios diferencian entre el capital social *bonding* (o lazos más fuertes y emocionales) y *bridging* (lazos más débiles e instrumentales) acorde a la clasificación de Putnam (2000), resaltando la importancia de las redes débiles (*weak ties*) y redes fuertes (*strong ties*) (Granovetter, 1973). En estos casos, la fortaleza de los vínculos depende de la frecuencia y cercanía de las relaciones sociales. Asimismo, en los estudios se emplean los niveles de análisis de redes sociales tanto a nivel sociocéntrico como egocéntrico, destacando las medidas de centralidad de grado y cercanía para calcular el capital social en el nivel de un solo actor, además de la medida de densidad en el nivel de toda la red. Además, la mayoría de estos estudios tienen sus implicaciones en el análisis de redes sociales respecto a la metodología utilizada. Los métodos cuantitativos utilizan cuestionarios de redes sociales (nivel sociocéntrico) (Wasserman y Faust, 1998; Crossley et al., 2015), y herramientas como el generador de nombres y posiciones (nivel egocéntrico) (Lin, 2001; Molina, 2005; Van Der Gaag y Snijders, 2005) para su posterior análisis. Por otra parte, los métodos cualitativos se basan en técnicas de mapeo visual como los círculos concéntricos (Kahn y Antonucci, 1980) y la realización de entrevistas y grupos de discusión. Por último, también encontramos estudios que utilizan métodos mixtos.

A continuación, se realiza una propuesta de clasificación de estos estudios en cuanto a la conceptualización y medición del capital social con el fin de contrastar las aportaciones de los artículos y observar si existen diferencias o patrones destacables.

## Conceptualización del Capital Social

La conceptualización del capital social se ha efectuado a partir de la propuesta de relación entre los componentes de redes, confianza, normas y sanciones (Baron, Field y Schuller, 2000; Halpern, 2005) y las dimensiones estructural, relacional y cognitiva (Nahapiet y Ghoshal, 1998) (tabla I).

En estos estudios existe un acuerdo en la concepción del capital social como los recursos que se desprenden de las relaciones sociales, por lo que en la mayoría de ellos el capital social se explica a través de aspectos estructurales, incidiendo en la posición en la red y en la cantidad de las relaciones. Asimismo, existen estudios que remarcan la parte más relacional atribuido a la confianza, enfatizando la calidad de las relaciones. Por último, encontramos estudios que reflejan la parte más cognitiva atribuida principalmente a normas comunes que tienen en cuenta el diálogo profesional y la identificación con el grupo. No obstante, la mayoría de estos estudios analizan el capital social como un concepto multidimensional en el que se trata de manera simultánea las diferentes dimensiones para obtener un análisis más completo del objeto de estudio desde diferentes aproximaciones.

**TABLA I.** Clasificación de la conceptualización del capital social

Artículo	CS ESTRUCTURAL	CS RELACIONAL	CS COGNITIVO	
	Redes	Confianza	Normas	Sanciones
Schiff, D., Herzog, L., Farley-Ripple, E., y Thum, L. (2015)	✓			
Colomer, S.E. (2014)	✓	✓	✓	✓
Daly, A., Moo-lenaar, N., Der-Martirosian, C., y Liou, Y. (2014)	✓	✓		
Liou, Y. y Daly, A. (2014)	✓	✓	✓	
Schreurs, B. y De Laet, M. (2014)	✓			
Ranieri, M., Manca, S., y Fini, A. (2012)	✓	✓		
Baker-Doyle, K. y Yoon, S. (2011)	✓			
Cannata, M. (2011)	✓	✓	✓	

Van Emmerik, H., Jawahar, I.M., Schreurs, B., y de Cuyper, N. (2011)	✓		✓	✓
Anderson, L. (2010)	✓	✓		
Frank, K. A. (2009)			✓	
Maier, A. y Youngs, P. (2009)	✓	✓	✓	
Penuel, W., Riel, M., Krause, A., y Frank, K. (2009)	✓	✓	✓	
Coburn, C. y Russell, J. (2008)	✓	✓		

Fuente: elaboración propia

## Medición del Capital Social

Para la medición del capital social se ha adoptado la propuesta de Lin (2001) y posterior adaptación de Carolan (2013) en el que se entiende como un activo en la red. Concretamente, en la tabla II se indican dos tipos de activos (recursos incrustados y posición de la red) y sus métodos de medición (recursos de las redes, recursos de los contactos, *brokerage*, *closure* y fuerza del vínculo).

Los recursos incrustados se centran en el valor de los recursos de las redes y de contactos a los que puede acceder un determinado actor en su red. Los recursos de las redes se refieren a aspectos como el rango, calidad, variedad y composición de las redes (consejo, información...), mientras que en los contactos se refieren a las características de los diferentes actores de la red en cuanto al status, ocupación y poder.

Por su parte, la posición de la red se refiere a la estructura de la red más allá del contenido. En esta aproximación existen diferentes dimensiones, según si se trata principalmente de *brokerage* o puente, la noción de *structural holes* y medidas de centralidad (resaltando la de intermediación) (Burt, 1992). En la dimensión *closure* se alude a aspectos generales de la red y de su cohesión, destacando la medida de densidad. Finalmente, la dimensión de fuerza del vínculo se refiere a la distinción entre redes fuertes y débiles (Granovetter, 1973), en cuanto a la intensidad, interacción y reciprocidad.

A continuación se muestran los resultados obtenidos según los diferentes estudios, donde se aprecia que la aproximación a los recursos incrustados aparece en prácticamente todos los estudios, tanto de redes como de contactos. Por otro lado, la posición de la red en referencia a su estructura suele tratarse en menor medida aunque de manera complementaria al contenido de los recursos.

TABLA II. Clasificación de la medición del capital social

Artículo	RECURSOS INCRUSTADOS		POSICIÓN DE LA RED		
	Recursos de las redes	Recursos de los contactos	Brokerage	Closure	Fuerza del vínculo
Schiff, D., Herzog, L., Farley-Ripple, E., y Thum, L. (2015)	✓			✓	
Colomer, S.E. (2014)		✓			✓
Daly, A., Moolenaar, N., Der-Martirosian, C., y Liou, Y. (2014)	✓	✓	✓		✓
Liou, Y. y Daly, A. (2014)	✓	✓	✓		
Schreurs, B. y De Laat, M. (2014)	✓		✓	✓	
Ranieri, M., Manca, S., y Fini, A. (2012)	✓	✓			✓
Baker-Doyle, K. y Yoon, S. (2011)	✓	✓	✓	✓	
Cannata, M. (2011)		✓			
Van Emmerik, H., Jawahar, I.M., Schreurs, B., y de Cuyper, N. (2011)		✓			
Anderson, L. (2010)	✓				✓
Frank, K. A. (2009)	✓	✓			
Maier, A. y Youngs, P. (2009)	✓	✓			
Penuel, W., Riel, M., Krause, A., y Frank, K. (2009)	✓	✓	✓	✓	
Coburn, C. y Russell, J. (2008)	✓				

Fuente: elaboración propia

## Conclusiones

En la presente revisión sistemática el capital social aparece como un concepto que demuestra la importancia de las interacciones y conexiones. Estos estudios tienen en común que la atención de los diversos retos se plantea desde un punto de vista relacional y considerando el rol de los vínculos sociales. De este modo, se adopta una perspectiva social centrada en el potencial de los lazos sociales que emerge como nueva propuesta a partir de las inversiones en las relaciones sociales, frente a las inversiones en habilidades o conocimientos que caracteriza el capital humano (Coleman, 1988). Por tanto, este concepto con su origen en la sociología (Bourdieu, 1986; Coleman, 1988) se convierte en un elemento clave en educación para entender los fenómenos complejos y contribuir así a la mejora del sistema educativo (Dika y Singh, 2002). Para lograr este propósito, el análisis de redes sociales se presenta como una herramienta a su servicio (Wasserman y Faust, 1998; Scott, 2000; Lin, 2001), teniendo en cuenta la multiplicidad de opciones de medidas y cálculos que ofrece esta técnica. Asimismo, se pueden detectar las redes más densas y las redes en las que se tienen que invertir más en las relaciones sociales para lograr una mayor cohesión. A su vez, se pueden detectar los actores clave y aislados para una más eficiente distribución y apropiación de los recursos.

Por otro lado, los 14 estudios de maestros analizados y correspondientes a los años 2008-2015, se centran principalmente en el potencial de las interacciones entre los maestros en redes formales e informales para movilizar recursos y para lograr posiciones en la red que les aporte beneficios para mejorar su red de apoyo, desarrollo profesional y éxito de sus estudiantes. En esta dirección, cada vez encontramos más estudios que pretenden demostrar la utilidad de atender a los lazos sociales entre los maestros (Dika y Singh, 2002; Smethem, 2007; Penuel et al., 2009; Daly y Finnigan, 2010; Moolenaar et al., 2010), hecho que implica destinar parte de la jornada laboral dentro los centros educativos para realizar discusiones, trabajos cooperativos, etc. Así pues, conviene a su vez centrar la atención en los actuales modelos emergentes de enseñanza que permiten crear y compartir conocimiento como el *co-teaching* o el *team teaching*. Del mismo modo, también se alude al concepto de las comunidades profesionales de aprendizaje para resaltar la importancia de la colectividad y del grupo en entornos profesionales. Estas

comunidades comparten y reflejan críticamente su práctica con el fin de mejorar la enseñanza de los estudiantes y profesores y se componen de interacciones colaborativas de educadores caracterizadas por diferentes elementos como unos valores y misiones compartidos, colaboración y confianza (Hord, 1997; Bolam, McMahon, Stoll, Thomas y Wallace, 2005). Para el fomento de estas interacciones, se deben tener en cuenta los elementos organizacionales y estructurales de los centros para proveer de los espacios físicos y temporales adecuados para la realización de seminarios, debates o talleres de formación, entre otros.

En estos estudios la confianza se conforma como un elemento básico del capital social (Coleman, 1990; Putnam, 1993; Fukuyama, 1995), para poder complementar los elementos más estructurales de la red en el terreno educativo y como un elemento esencial en las escuelas (Bryk y Schneider, 2002; Tschannen-Moran, 2014). Sin embargo, el papel de la confianza en la conceptualización y medida del capital social no se encuentra en la literatura claramente definido, debido a la diversidad de perspectivas que los teóricos reflejan sobre el concepto (Coleman, 1988; Fukuyama, 1995; Putnam, 1995; Baron, Field y Schuller, 2000; Halpern, 2005). Por tanto, se necesitarían más estudios específicos donde se explore la contribución de la confianza en los elementos estructurales y cognitivos, así como en el papel que juega en la distribución de recursos y las diferentes posiciones en la red. En esta revisión, el papel de la confianza suele atribuirse a parte de la definición conceptual del capital social, en su dimensión más relacional, y en referencia a redes más instrumentales como la de pedir consejo, ayuda o conocimiento, como fuente de apoyo y soporte para los maestros. En este sentido, la red de consejo es la que más se utiliza, más allá del uso de redes más expresivas (Lin, 1982; Ibarra, 1993).

Teniendo en cuenta la clasificación del capital social como un concepto multidimensional (Nahapiet y Ghoshal, 1998), a partir de los resultados se detecta la necesidad de más estudios que exploren la parte cognitiva en referencia a los valores, significados y conocimientos compartidos que comprendan la cultura y el clima de la escuela como condiciones necesarias para facilitar las interacciones sociales y promover el capital social. En esta dirección, crear un clima colaborativo y de apoyo entre los profesionales es un elemento fundamental para incrementar el capital social y mejorar el desarrollo profesional de los maestros (Bryk y Schneider, 2002; Cosner, 2009; Daly, Moolenaar, Der-Martirosian y

Liou, 2014). Para ello, se debe fomentar la reflexión compartida entre los profesionales de los centros, así como la concreción de un objetivo en común. Del mismo modo, se debe contemplar la dimensión estructural de redes y relacional de confianza, puesto que todos estos aspectos de manera integrada pueden contribuir a la mejora del nivel de colaboración de los individuos dentro de la red (Nahapiet y Ghoshal, 1998; Goddard, 2003; Liou y Daly, 2014).

Asimismo, a través de la medición del capital social se enfatiza la importancia de los recursos valiosos en las relaciones de los maestros, ya sea a través de las redes o de los contactos, mostrando la necesidad de más estudios que resalten la posición en la red de los actores, la cohesión y fortaleza de los vínculos. Respecto a la posición de la red, las medidas de centralidad son un elemento clave a considerar para relacionar la posición de maestros en la escuela, o durante su formación en la universidad, con su desarrollo académico y profesional (Hommes et al., 2012; Liou et al., 2016). Por otra parte, son relevantes las medidas de cohesión para valorar como una red más densa y conectada permite a los maestros trabajar de manera colaborativa para mejorar sus habilidades y fomentar el intercambio de recursos (Daly, 2010; Cívís et al., 2017).

Finalmente, la mayoría de los estudios de maestros tratan sobre maestros con experiencia, mientras que tan sólo se incluye un caso en el que se trate de futuros maestros. En la literatura existen estudios que demuestran la importancia de adoptar esta perspectiva social en la formación de los maestros y en los programas universitarios (Liou et al., 2016; Cívís et al., 2017), demostrando que además un clima colaborativo promueve a su mejora del desarrollo profesional (Bryk y Schneider, 2002; Cosner, 2009; Daly et al., 2014). Por otro lado, solamente un caso alude a los maestros noveles que se encuentran en el inicio de su carrera profesional, teniendo presente que es en este momento en el que precisan más apoyo de sus compañeros para poderse desarrollar profesionalmente (Ewing y Manuel, 2005; Gasevic, Zouaq y Janzen, 2013). Diversos estudios demuestran que proveer de soporte efectivo a los maestros noveles puede contribuir a mejorar su retención (Yuen, 2012; De Vries, Vand de Grift y Jansen, 2013) Por eso, es especialmente importante entender la motivación y experiencia de los maestros en sus primeros años de profesión (Smethem, 2007) con tal de poder encaminar las adecuadas políticas para evitar el abandono de sus carreras profesionales.

En resumen, este estudio muestra el estado de la cuestión de las investigaciones sobre maestros desde una perspectiva social y relacional. Así se constata cómo empiezan a existir diversos estudios sobre la profesión que complementan y ofrecen otra perspectiva a los estudios que destacan los aspectos más curriculares, evaluativos y de planificación (Nelson y Sassi, 2000; Yuen, 2012; Nilsson y Loughran, 2012). Sin embargo, la muestra de 14 estudios pone de manifiesto que se trata de una línea de trabajo aún incipiente.

Además, a partir de esta revisión sistemática se observa cómo, a pesar de que se enfatiza la importancia de atender al capital social y a las redes sociales, no existe un acuerdo generalizado sobre cómo analizar y explorar el capital social de los maestros, puesto que se adoptan diferentes posturas al respecto, a partir de la utilización de distintas aproximaciones de comprensión y medición del concepto. Por tanto, en este campo aún faltaría por explorar en más detalle estos aspectos para comprender el funcionamiento y la utilidad del capital social y llegar a un mayor consenso. Para ello, los esfuerzos se deben dirigir en entender el potencial del capital social en educación de manera general y en los maestros de manera específica, para llegar a un acuerdo común de análisis y evaluación y poder seguir así evolucionando en esta dirección, tanto en el ámbito de la formación inicial como de la formación continuada y su ejercicio profesional.

Toda investigación tiene una serie de limitaciones. En primer lugar, siguiendo la estrategia de búsqueda, los estudios de la literatura que se han revisado son los que se encuentran dentro de las bases de datos seleccionadas por su relevancia y alcance internacional, teniendo presente que pueden existir otros estudios del mismo ámbito no contemplados para este estudio. Pese a ello, la inmensa cantidad de estudios incluidos se han revisado de manera exhaustiva, teniendo presente la dificultad de gestionar el gran abasto de información y proceder a su clasificación.

Por último, de cara a la prospectiva del estudio, debido a la novedad de los temas tratados y a su auge exponencial, en un futuro cercano sería conveniente realizar el seguimiento y análisis de los nuevos estudios que pueden aportar nuevos y relevantes hallazgos en este terreno. Del mismo modo, sería interesante formalizar una propuesta de un nuevo modelo para poder analizar el capital social en contextos educativos que recoja y sintetice las diversas aproximaciones exploradas en cuanto a su conceptualización y medición.

## Referencias bibliográficas

- Anderson, L. (2010). Embedded, emboldened, and (net)working for change: Support-seeking and teacher agency in urban, high-needs schools. *Harvard Education Review*, 80(4), 541-572.
- Baker-Doyle, K. J., y Yoon, S. A. (2011). In search of practitioner-based social capital: A social network analysis tool for understanding and facilitating teacher collaboration in a US-based STEM professional development program. *Professional Development in Education*, 37(1), 75-93.
- Baron, S., Field, J., y Shuller T. (2000). *Social Capital. Critical Perspectives*. Nueva York: Oxford University Press.
- Bauman, Z. (1999). *Modernidad líquida*. Buenos Aires, Argentina: Fondo de Cultura Económica.
- Baumard, P., y Starbuck, W. H. (2005). Learning from failures: Why it may not happen. *Long Range Planning*, 38, 281-298
- Bolam, R., McMahon, A., Stoll, L., Thomas, S., y Wallace, M. (2005). *Creating and sustaining professional learning communities*. London: General Teaching Council for England, Department for Education and Skills.
- Bolívar, A. (2006). Familia y escuela: dos mundos llamados a trabajar en común. *Revista de Educación*, 339, 119-146.
- Borgatti, S.P., Everett, M.G. y Johnson, J.C. (2013). *Analyzing Social Networks*. London: Sage Publications.
- Borgatti, S. P., y Foster, P. (2003). The network paradigm in organizational research: A review and typology. *Journal of Management*, 29(6), 991-1013.
- Bourdieu, P. (1986). *Handbook of theory and research for the sociology of education*. New York: Greenwood Press.
- Bryk, A. S., y Schneider, B. (2002). *Trust in schools: A core resource for improvement*. New York, NY: Russell Sage Foundation.
- Burt, R. S. (1992). *Structural holes*. Cambridge, MA: Harvard University Press.
- Burt, R. S. (2004). Structural holes and good ideas. *American Journal of Sociology*, 110, 349-399.
- Cannata, M. (2011). The role of social networks in the teacher job search process. *The Elementary School Journal*, 111(3), 477-500.

- Castells, M. (1996). *Rise of the network society*. Cambridge, MA: Blackwell.
- Carolan, B. V. (2013). *Social network analysis and education. Theory, methods, and applications*. Thousand Oaks, CA: Sage Publications.
- Civís, M., López, S., y Díaz-Gibson, J. (2017). El capital social dels estudiants del darrer curs de Magisteri. Una aproximació a les seves xarxes socials en el marc de la facultat. *REIRE Revista d'Innovació i Recerca en Educació*, 10(1), 34-56.
- Clayton, C. D. (2007). Curriculum Making as Novice Professional Development: Practical Risk Taking as Learning in High-Stakes Times. *Journal of Teacher Education*, 58(3), 216-230.
- Coburn, C. E., y Russell, J. L. (2008). District policy and teachers' social networks. *Education Evaluation and Policy Analysis*, 30(3), 203-235.
- Coleman, J. (1988). Social capital in the creation of human capital. *American Journal of Sociology*, 94, S95-S120.
- Coleman, J. S. (1990). *Foundations of social theory*. Cambridge, MA: Harvard University Press.
- Colomer, S. E. (2014). Latina Spanish High School Teachers' Negotiation of Capital in New Latino Communities. *Bilingual Research Journal*, 37(3), 349-365.
- Cosner, S. (2009). Building organizational capacity through trust. *Educational Administration Quarterly*, 45(2), 248-291.
- Cross, R., Borgatti, S., & Parker, A. (2002). Making invisible work visible. *California Management Review*, 44(2), 25-46
- Crossley, N., Bellotti, E., Edwards, G., Everett, M.G., Koskinen, J., y Tranmer, M. (2015). *Social network analysis for ego-nets*. London, UK: SAGE.
- Daly, A. J. (2010). *Social network theory and educational change*. Cambridge, MA: Harvard Education Press.
- Daly, A. J., y Finnigan, K. (2010). A bridge between worlds: understanding network structure to understand change strategy. *Journal of Educational Change*, 11(2), 111-138.
- Daly, A. J., y Finnigan, K. S. (2011). The ebb and flow of social network ties between district leaders under high-stakes accountability. *American Educational Research Journal*, 48(1), 39-79.
- Daly, A. J., Moolenaar, N., Der-Martirosian, C., y Liou, Y.-H. (2014). Accessing capital resources: Investigating the effects of teacher human and social capital on student achievement. *Teachers College Record*, 116(7), 1-42.

- Darling-Hammond, L. (2006). Constructing 21st-Century Teacher Education. *Journal of Teacher Education*, 57(3), 300-314.
- De Vries, S., Van de Grift, W. J., y Jansen, E. P. (2013). Teachers' beliefs and continuing professional development. *Journal of Educational Administration*, 51(2), 213-231.
- Dika, S. y Singh K. (2002). Applications of Social Capital in Educational Literature: A Critical Synthesis. *Review of Educational Research*, 72 (1), 31-60.
- Ewing, R., y Manuel, J. (2005). Retaining quality early career teachers in the profession. *Change: Transformations In Education*, 8(1), 1-16.
- Frank, K. A. (2009). Quasi-ties: Directing resources to members of a collective. *American Behavioral Scientist*, 52, 1613-1645.
- Freeman, L.C (1979). Centrality in Social Networks Conceptual Clarification. *Social Networks*, 1, 215-239.
- Fukuyama, F. (1995). *Trust: The social virtues and the creation of prosperity*. New York: Free Press.
- Fukuyama, F. (2001). Social Capital, civil society and development. *Third World Quarterly*, 22 (1), 7-20.
- Gasevic, D., Zouaq, A., y Janzen, R. (2013). "Choose your classmates, your GPA is at stake!": The association of cross-class social ties and academic performance. *American Behavioral Scientist*, 57(10), 1460-1479.
- Goddard, R. (2003). Relational networks, social trust, and norms: A social capital perspective on students' chances of academic success. *Educational Evaluation and Policy Analysis*, 25(1), 59-74.
- Granovetter, M. S. (1973). The strength of weak ties. *American Journal of Sociology*, 78 (6), 1360-1380.
- Halpern, D. (2005). *Social Capital*. Cambridge: Polity Press.
- Hâncean, M-G., Molina, J.L., y Lubbers, M. (2016). Recent advancements, developments and applications of personal network analysis. *International Review of Social Research*, 6(4), 137-145.
- Hawkins, R. L., y Maurer, K. (2010). Bonding, bridging and linking: How social capital operated in New Orleans following Hurricane Katrina. *British Journal of Social Work*, 40(6), 1777-1793.
- Hommel, J., Rienties, B., Grave, W. D., Bos, G., Schuwirth, L., & Scherpbier, A. (2012). Visualising the invisible: A network approach to reveal the informal social side of student learning. *Advances in Health Education*, 17(5), 743-757.

- Hord, S. (1997). *Professional learning communities: Communities of continuous inquiry and improvement*. Austin, TX: Southwest Educational Development Laboratory.
- Ibarra, H. (1993). Personal networks of women and minorities in management: a conceptual framework. *Academy of Management Review*, 18(1) 58-67.
- Kahn, R. L., y Antonucci, T. C. (1980). Convoys over the life course: Attachment, roles and social support. En P. B. Baltes y O. G. Brim (Eds.). *Life-span development and behaviour* (pp. 383-405). New York: Academic Press.
- Kilduff, M., y Tsai, W. (2003). *Social Networks and Organizations*. London: Sage Publications.
- Liberati, A., Altman, D.G., Tetzlaff, J., Mulrow, C., Gøtzsche, P.C., Ioannidis, J.P.A., et al. (2009). The PRISMA Statement for Reporting Systematic Reviews and Meta-Analyses of Studies That Evaluate Health Care Interventions: Explanation and Elaboration. *PLoS Medicine*, 6(7), e1000100.
- Lin, N. (1982) Social Resources and Instrumental Action. In: P. V. Marsden & N. Lin (Eds). *Social Structure and Network Analysis* (pp. 131-45). Beverly Hills, CA: Sage.
- Lin, N. (1999). Building a network theory of social capital. *Connections*, 22(1), 28-51
- Lin, N. (2001). *Social capital: A theory of social structure and action*. (1st Ed.). New York: Cambridge University Press.
- Liou, Y.-H., y Daly, A. J. (2014). Closer to Learning. Social Networks, Trust, and Professional Communities. *Journal of School Leadership Volume*, 24, 753-795.
- Liou, Y.-H., Daly, A. J., Canrinus, E. T., Forbes, C. A., Moolenaar, N. M., Cornelissen, F., y Hsiao, J. (2016). Mapping the social side of pre-service teachers: connecting closeness, trust, and efficacy with performance. *Teachers and Teaching*, 1-23.
- Maier, A., y Youngs, P. (2009). Teacher preparation programs and teacher labor markets: How social capital may help explain teachers' career choices. *Journal of Teacher Education*, 60(4), 393-407.
- Marquès, P. (1999). Criterios para la clasificación y evaluación de espacios web de interés educativo. *Educar*, 25, 95-111.
- McVeigh, M. E., y Bates, M. J. (2010). Citation Indexes and the Web of Science. *Encyclopedia of Library and Information Science*, 2.

- Miller, P. (2012). Community-Based Education and Social Capital in an Urban After-School Program. *Education and Urban Society*, 44(1), 35-60.
- Molina, J.L. (2005). El estudio de las redes personales: contribuciones, métodos y perspectivas. *Revista de Metodología de las Ciencias Sociales*, (10), 71-105.
- Morin, E. (2004). *Introducción al pensamiento complejo*. Barcelona: Gedisa.
- Moolenaar, N. M. (2010). *Ties with potential. Nature, antecedents, and consequences of social networks in school teams*. Unpublished doctoral dissertation. University of Amsterdam, The Netherlands.
- Moolenaar, N., Daly, J., y Slegers, P. (2010). Occupying the principal position: examining relationships between transformational leadership, social network position, and schools' innovative climate. *Educational Administration Quarterly*, 46(5), 623-670.
- Nahapiet, J., y Ghoshal, S. (1998). Social capital, intellectual capital and the organizational advantage. *Academy of Management Review*, 23(2), 242-266.
- Nelson, B. S. y Sassi, A. (2000). Shifting approaches to supervision: The case of mathematics supervision. *Educational Administration Quarterly*, 36(4), 553-584.
- Nilsson, P. y Loughran, J. (2012). Exploring the development of pre-service elementary teachers' pedagogical content knowledge, *Journal of Science Teacher Education*, 23(7), 699-721.
- Penuel, W. R., Riel, M. R., Krause, A., y Frank, K. A. (2009) Analyzing teachers' professional interactions in a school as social capital: A social network approach. *Teachers College Record*, 111(1), 124-163.
- Penuel, W. R., Sun, M., Frank, K. A., y Gallagher, H. A. (2012). Using Social Network Analysis to Study How Collegial Interactions Can Augment Teacher Learning from External Professional Development. *American Journal of Education*, 119(1), 103-136.
- Piqué, B., Comas, A., y Lorenzo, N. (2010). *Estratègies de pràctica reflexiva en la formació inicial de mestres d'educació infantil*. Barcelona: Graó.
- Portes, A. (1998). Social capital. *Annual Review of Sociology*, 24, 1-24.
- Putnam, R. D. (1993). *The prosperous community: Social capital and public life*. *American Prospect*, 4, 35-42.

- Putnam, R. D. (1995). Bowling alone: America's declining social capital. *Journal of Democracy*, 6 (1), 65-78.
- Putnam, R. D. (2000). *Bowling alone: The collapse and revival of American community*. New York: Simon & Schuster.
- Ranieri, M., Manca, S., y Fini, A. (2012). Why (and how) do teachers engage in social networks? An exploratory study of professional use of Facebook and its implications for lifelong learning. *British Journal of Educational Technology*, 43(5), 754-769.
- Rienties, B., y Kinchin, I. M. (2014). Understanding (in)formal learning in an academic development programme: A social network perspective. *Teaching and Teacher Education*, 39, 123-135.
- Schiff, D., Herzog, L., Farley-Ripple, E., y Thum, L. (2015). Teacher networks in Philadelphia: Landscape, engagement, and value. *Perspectives on Urban Education*, 12 (1), 1-17.
- Schleicher, A. (2012). *Preparing Teachers and Developing School Leaders for the 21st Century: Lessons from around the World*. Paris: OECD Publishing.
- Schreurs, B., y De Laat, M. (2014). The Network Awareness Tool: A web 2.0 tool to visualize informal networked learning in organizations. *Computers in Human Behavior*, 37, 385-394.
- Scott, J. (2000). *Social network analysis*. (2nd ed.). London: Sage.
- Slegers, P. J. C., Den Brok, P., Verbiest, E., Moolenaar, N. M., & Daly, A. J. (2013). Towards conceptual clarity: A multidimensional, multilevel model of professional learning communities in Dutch elementary schools. *Elementary School Journal*, 114(1), 118-137.
- Smethem, L. (2007). Retention and intention in teaching careers: will the new generation stay? *Teachers and Teaching*, 13 (5), 465-480.
- Tschannen-Moran, M. (2014). *Trust Matters: Leadership for Successful Schools*. (2nd ed.). San Francisco: Jossey-Bass.
- Uzzi, B. (1997). Social structure and competition in interfirm networks: the paradox of embeddedness. *Administrative Science Quarterly*, 42 (1).
- Vaillant, E. y Manso, J. (2013). Teacher education programs: learning for worldwide inspiring experiences. *Journal of Supranational Policies of Education*, (1), 95-115.
- Van Der Gaag, M., y Snijders, T. (2005). The resource generator: social capital quantification with concrete items. *Social Network*, 27(1), 1-29.

- Van Emmerik, H., Jawahar, I.M., Schreurs, B., y de Cuyper, N. (2011). Social capital, team efficacy and team potency: The mediating role of team learning behaviors. *Career Development International*, 16(1), 82-99.
- Wasserman, S., y Faust, K. (1998). *Social network analysis: Methods and applications*. New York: Cambridge University Press.
- Woolcock, M. (2001). The place of social capital in Understanding Social and Economic Outcomes. *SUMA Canadian Journal of Policy Research*, 2(10),11-17.
- Yuen, L. H. (2012). The impact of continuing professional development on a novice teacher. *Teacher Development*, 16(3), 387-398.

**Información de contacto:** Susana López Solé. Universidad Ramón Llull, Facultad de Psicología, Ciencias de la Educación y del Deporte, Educación, c/ del Císter, 34, 08022 Barcelona. E-mail: susanals@blanquerna.url.edu



# La enseñanza de las ciencias basada en indagación. Una revisión sistemática de la producción española

## Inquiry-based Science Education. A systematic review of Spanish production

DOI: 10.4438/1988-592X-RE-2017-381-388

David Aguilera Morales  
Tobías Martín-Páez  
Víctor Valdivia-Rodríguez  
Ángela Ruiz-Delgado  
Leticia Williams-Pinto  
José Miguel Vilchez-González  
Francisco Javier Perales-Palacios

*Universidad de Granada*

### Resumen

Ante la creciente relevancia otorgada a la enseñanza basada en indagación tanto en el ámbito internacional como nacional, este artículo pretende aunar aquellos trabajos referentes a esta línea de investigación publicados entre 2007-2017 en las revistas españolas del área de Didáctica de las Ciencias Experimentales. Se han revisado 1941 artículos, incluyéndose 55 de ellos tras la aplicación de los criterios de selección establecidos, todo ello según las directrices de la Declaración PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses). Los resultados muestran cuáles son los autores más prolíficos, la evolución de la producción científica, las etapas educativas consideradas, los tipos de estudios, sus diseños y las implicaciones educativas de su aplicación en las aulas. A la luz de estos, podemos concluir que no abundan autores o autoras con una excelsa producción referente a esta temática, pues esta línea de investigación es aún incipiente en España. Además, la atención dedicada a las distintas etapas educativas por la investigación en Didáctica de

las Ciencias Experimentales no es equitativa, siendo las grandes perjudicadas las etapas de Educación Infantil y Primaria. Existen más artículos de naturaleza teórica que empírica, dado que queda patente la falta de estudios experimentales que testen este enfoque de enseñanza. No obstante, las implicaciones educativas derivadas de las publicaciones españolas son similares a las internacionales. Finalmente, hemos de abogar por un incremento de los estudios dirigidos a evaluar los beneficios de la enseñanza basada en indagación en las diferentes etapas educativas, incidiendo en las etapas de Educación Infantil y Primaria.

*Palabras clave:* Enseñanza de las ciencias; enseñanza basada en indagación; implicaciones educativas; revistas españolas; revisión sistemática.

### **Abstract**

Due to the growing importance given to teaching based on inquiry both at the international and national level, this article aims to further work related to this line of research published between 2007 and 2017 in the Spanish area of Didactic of Experimental Sciences. We reviewed 1941 articles, including 55 criteria for the selection, all according to the guidelines of the PRISMA statement (Preferred Reporting Items for Systematic Reviews and Meta-Analyses). The results show the most prolific authors, the evolution of scientific production, the educational stages considered, the types of studies, the designs and the educational implications of their application in the classrooms. In light of these, we can conclude that there are not many authors with an excellent production regarding this subject, because this research's line is still incipient in Spain. The attention devoted to the different educational stages by research in Didactic of Experimental Sciences is not equitable, with the most affected being the stages of Infantile and Primary Education. There are more articles of a theoretical nature than empirical ones, given the lack of experimental studies that prove this teaching approach is patent. However, the educational implications derived from Spanish publications are similar to the international ones. Finally, we have to advocate for an increase in the number of studies aimed at evaluating the benefits of teaching based on inquiry in the different educational stages, focusing on the stages of Pre-school Education and Primary Education.

*Keywords:* Science education; teaching based on inquiry; educational implications; Spanish journals; systematic review.

## **Introducción**

La indagación como estrategia de enseñanza y aprendizaje de las ciencias está abriéndose paso como una alternativa sugerente ante la necesidad

de mejorar la actual forma de enseñar ciencias en los centros educativos. Algunos de los objetivos que se persigue con ello serían los siguientes:

- (1) Mejorar la imagen y la actitud hacia la ciencia.
- (2) Contribuir a la alfabetización científica.
- (3) Incentivar las vocaciones científicas.

En ello sin duda ha influido la recomendación de su uso por parte del Informe Rocard (2007), y más recientemente la publicación del informe de la Comisión Europea: «Science Education for Responsible Citizenship» (Hazelkorn, 2015), en el que se abunda en dicha recomendación. Así, se señala que «Education policies and systems should: ... Support schools, teachers, teacher educators and students of all ages to adopt an inquiry approach to science education as part of the core framework of science education for all...» (p. 10).

A pesar de la importancia otorgada en los últimos tiempos a esta línea de investigación, existe un número reducido de trabajos de síntesis sobre esta temática. Concretamente, en la última década, encontramos cuatro estudios de revisión sobre la enseñanza basada en indagación en el panorama internacional (Minner, Levy y Century, 2010; Alfieri, Brooks, Aldrich y Tenenbaum, 2011; Furtak, Seidel, Iverson y Briggs, 2012; Lazonder y Harmsen, 2016), mientras que en el ámbito nacional solo disponemos del trabajo de Romero-Ariza (2017). Aunque las diferencias entre la revisión de Ro-Ariza (2017) y el aquí expuesto son evidentes, dado que la autora realiza una revisión narrativa en la que comenta resultados obtenidos por otros estudios de revisión, en este caso meta-análisis, que versan sobre la temática abordada.

Para ampliar la información al respecto, en este artículo se revisan las investigaciones publicadas entre los años 2007 y 2017 en las revistas españolas específicas de Didáctica de las Ciencias Experimentales (DCE) a partir de las siguientes preguntas de investigación:

- (1) ¿Quiénes son los autores más prolíficos en esta línea de investigación? ¿Qué grado de colaboración existe?
- (2) ¿Cuántos artículos abordan la enseñanza basada en indagación? ¿Cuál ha sido su evolución en el tiempo? ¿En qué países se han llevado a cabo?
- (3) ¿En qué etapas educativas se está investigando?

- (4) ¿Qué tipo de estudios se publican sobre esta temática? ¿Qué diseños tienen?
- (5) ¿Qué implicaciones educativas exponen en relación a la enseñanza basada en indagación?

## Indagación y enseñanza de las ciencias

Lo primero que habría que definir en cualquier trabajo sobre indagación es en qué sentido se habla de ella. Como señala Couso (2014), el término «indagación» (inquiry, en inglés) es «sorprendentemente polisémico en la literatura educativa». Así, al hablar de indagación podemos hacer referencia a (Barrow, 2006, citado por Couso, 2014):

- La capacidad de investigar científicamente, si centramos la atención en el desarrollo de esta capacidad cognitiva.
- La naturaleza de la investigación científica, si focalizamos en que el alumnado entienda la metodología científica.
- Las estrategias de enseñanza-aprendizaje que persiguen que el alumnado, además de comprender conceptos científicos, adquiera capacidades «de» y «sobre» la indagación científica.

Las dos primeras acepciones hacen referencia a contenidos a enseñar, de procedimientos científicos en el primer caso y sobre ciencia en el segundo; la tercera, por el contrario, es una forma de llevar a cabo la indagación. Se habla, en este último caso, de enseñanza de las ciencias por / basada en / centrada en indagación (IBSE, por sus siglas en inglés Inquiry-Based Science Education), o de aprendizaje por / basado en indagación (IBL, Inquiry-Based Learning).

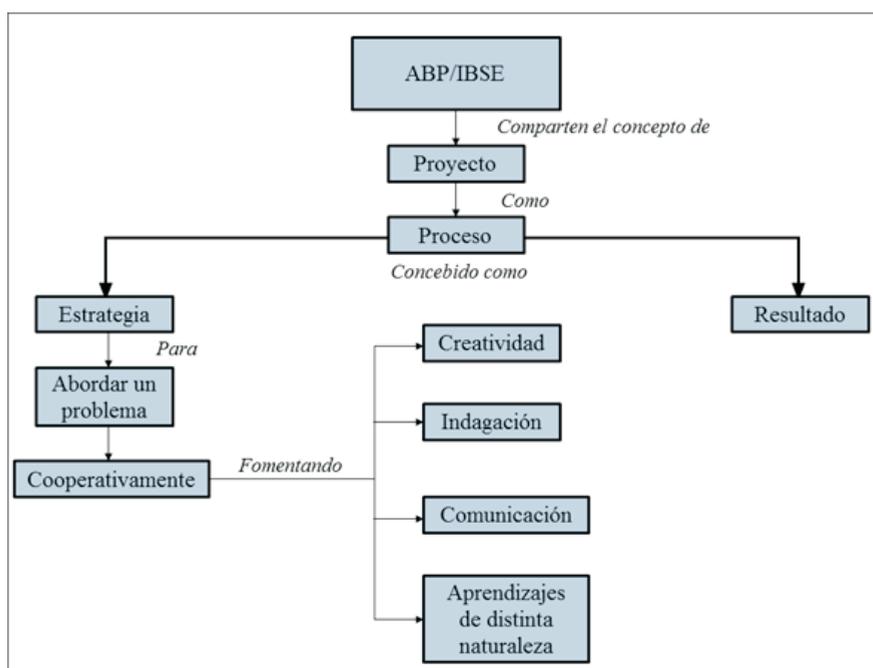
No obstante, se hace precisa previamente una clarificación conceptual para delimitar esta estrategia y diferenciarla de otras procedentes de diversos ámbitos de las ciencias sociales y de la salud, con algunos elementos comunes, tales como la enseñanza por proyectos o el aprendizaje basado en problemas (ABP, en ambos casos). En la figura I hemos tratado de facilitar tal distinción, exponiéndose a continuación algunas *definiciones* sobre indagación<sup>1</sup>:

---

<sup>(1)</sup> Los siguientes párrafos están adaptados de Martínez-Aznar (2017).

- «Es el proceso intencional de identificación y análisis de problemas, la distinción de alternativas y la evaluación de experiencias para resolverlos, la planificación de investigaciones, la construcción de modelos, la investigación sobre hipótesis, la búsqueda de información, el debate entre compañeros y la construcción de argumentos coherentes» (Linn, Davis y Bell, 2004, p. 4).
- «Comprender las distintas formas en que los científicos realizan su trabajo, valorar la potencialidad de la observaciones, habilidad para formular preguntas investigables y emitir hipótesis, utilizar distintos tipos de datos para buscar patrones y confirmar o rechazar las predicciones, construir y defender modelos y argumentos, juzgar explicaciones alternativas y lograr una mejor comprensión del carácter provisional y evolutivo de la ciencia, y su origen en la actividad humana, el contexto y la cultura en que se desarrolla y utiliza» (Crawford, 2007, p. 614).

FIGURA I. Mapa conceptual. Fuente: adaptado de Perales y Ayerbe (2016).



La IBSE, en la que centramos nuestra atención a partir de ahora, se suele proponer como alternativa a la enseñanza tradicional. Sus características principales son (Couso, 2014):

- Se organizan en entornos de investigación, normalmente de tipo práctico.
- El alumnado adquiere un papel más activo que en entornos tradicionales, y el profesorado, más pasivo (actúa de guía).
- Se da mucha importancia a la actitud y a la motivación.
- Los alumnos, que en general trabajan en grupo, se plantean preguntas y obtienen datos (o usan datos disponibles).
- El alumnado goza de mucha autonomía y capacidad de decisión.
- Se diseñan en etapas que simulan la metodología científica.

En función del grado de autonomía del alumnado, Windschitl (2003) diferencia desde las propuestas centradas en la *confirmación de experiencias* previamente discutidas en el aula hasta aquellas que se plantean como *indagación abierta* (el alumnado decide sobre qué investigar y cómo hacerlo), pasando por estrategias de *indagación estructurada* (el profesor proporciona tanto la pregunta como el procedimiento para resolverla) y de *indagación guiada* (el docente proporciona la pregunta y los alumnos deciden cómo resolverla), incidiendo cada una de ellas en mayor o menor medida en distintas destrezas científicas del alumnado (Tabla I).

En el ámbito español podemos encontrar precedentes de esta línea investigación, concretamente en los trabajos de D. Gil y colaboradores, así como de la red IRES (Investigación y Renovación Escolar). En el primer caso, centrado específicamente en la enseñanza y aprendizaje de las ciencias, Gil (1990, 1993) apuesta por el modelo de resolución de problemas por investigación, estableciendo una analogía entre la investigación científica y la resolución de problemas, donde el investigador novel (estudiante) trabaja bajo la dirección y supervisión del investigador principal (profesor), utilizando al efecto problemas abiertos (preferentemente cualitativos). En el segundo caso, la propuesta tiene un enfoque más transversal, considerando la investigación como un principio didáctico integrador (Cañal, 1987) que debiera impulsar en el aprendiz el espíritu científico y las habilidades propias de la metodología científica; ello debería impregnar los elementos más característicos del currículo escolar como los objetivos, contenidos, relaciones de comunicación, etc.

**TABLA I.** Modelos de indagación y destrezas desempeñadas por el alumnado. Fuente: Bevins y Price (2016).

<b>Modelos de indagación</b>	<b>Destrezas de indagación agrupadas en dimensiones</b>				
<b>Nivel</b>	<i>1. Plantear preguntas científicamente</i>	<i>2. Basarse en pruebas</i>	<i>3. Explicar a partir de pruebas</i>	<i>4. Explicar a partir de conocimientos</i>	<i>5. Comunicar y justificar</i>
<b>3. Abierta</b>	Formulan sus preguntas	Determinan aquello que constituye una prueba y lo recaban	Formulan explicaciones después de agrupar sus pruebas	Examinan otros recursos para formular explicaciones sobre su trabajo	Buscan cómo formular argumentos lógicos para comunicar sus explicaciones
<b>2. Guiada</b>	Seleccionan preguntas y plantean otras nuevas	Se les orienta para que recaben ciertos datos	Se les guía en el proceso de formular explicaciones a partir de pruebas	Se les dirige hacia áreas y fuentes de conocimiento científico	Se les orienta hacia estrategias para desarrollar la comunicación
<b>1. Estructurada</b>	Clarifican o concretan las preguntas que se les asignan	Reciben datos y se les pide su análisis	Se les proporcionan ejemplos de cómo usar las pruebas para formular explicaciones	Se proporcionan posibles conexiones con el conocimiento científico	Se les proporcionan directrices a usar para enfocar la comunicación
<b>0. Confirmación / actividades de verificación</b>	Dan respuesta a las preguntas proporcionadas por otros	Reciben datos y se les explica cómo analizarlos	Reciben explicaciones sobre sus resultados	Se les indica la conexión precisa entre los conocimientos	Reciben pasos y procedimientos específicos para la comunicación

## Indagación en el aula

En relación con su implementación en el aula, las fases de una propuesta IBSE son, en consonancia con la investigación científica, las siguientes (Martínez-Chico, 2013):

- (1) Identificar problemas o cuestiones de carácter científico (cuyas respuestas se puedan confirmar o rechazar usando pruebas).
- (2) Emitir hipótesis como posibles respuestas al problema o cuestión (explicaciones justificadas).
- (3) Buscar pruebas que confirmen o refuten la hipótesis (mediante experimentos o búsqueda de información).

- (4) Analizar e interpretar resultados.
- (5) Extraer conclusiones y comunicarlas.

Aunque la IBSE admite diversas interpretaciones (Romero-Ariza, 2017) que dependen de la propia concepción de indagación y de las experiencias docentes previas (Windschitl, 2003), hecho este que dificulta la investigación e incluso la discusión sobre indagación (Yeomans, 2011), todas coinciden en dos aspectos (Couso, 2014):

- Hay que llevar al aula la autenticidad de la metodología científica.
- Hay que involucrar y motivar a los alumnos.

Todo ello dibuja un escenario en el que las cuestiones a tener en cuenta y a ejecutar son complejas y, por tanto, difíciles de abordar, postulándose esta situación como una de las razones por las que las estrategias basadas en indagación no gozan aún de mucha presencia en las aulas de ciencias esgrimiéndose otras como: la falta de preparación del profesorado (Forbes y David, 2010; Lucero, Valcke y Schelles, 2013), la escasez de recursos, limitación temporal, currículo extenso y presión de los padres (Anderson, 1996) o la propia concepción de indagación y el modo de implementarla en el aula (Cuevas, Lee, Hart y Deakort, 2005; McDonald y Butler Soger, 2008).

A ello habría que añadir que los maestros en ejercicio y en formación disponen, en general, de un conocimiento científico fragmentado, superficial y poco sólido (Murphy et al., 2007), además de un cierto rechazo hacia las ciencias y bajos niveles de confianza (Vázquez y Manassero, 2008). Así, en particular, algunas limitaciones detectadas en maestros en formación para la implementación de actividades de indagación se relacionan con la elección de la pregunta a investigar, el diseño y la ejecución de la investigación (Lucero et al., 2013; Crujeiras y Puig, 2014).

Una vez establecido lo que en este trabajo se entiende por indagación (IBSE), sus características principales y algunos de los principales inconvenientes para que estas estrategias alcancen las aulas, quedaría por estudiar si las que se han implementado han alcanzado el impacto deseado. Aunque hay autores que consideran que la indagación es el mejor método para enseñar y aprender ciencias (Bevins y Price, 2016), o que «las estrategias de enseñanza que comprometen al estudiante activamente en el proceso de aprendizaje a través de investigaciones

científicas, incrementan la comprensión conceptual más que las estrategias que se basan en técnicas pasivas» (Minner, Levy y Century, 2010, p. 493), también los hay que ponen de manifiesto que no es así (Cobern et al., 2010). De hecho, existen trabajos que correlacionan negativamente las metodologías de indagación en el aula y el rendimiento académico en las pruebas PISA (Areepattamannil, 2012; McConney, Oliver, Woods-McConney, Schibeci y Maor, 2014, citados por Romero-Ariza, 2017).

A este respecto, cabe destacar que, como se ha descrito, son pocos los casos a analizar. No obstante, sí se pueden comentar algunos resultados. En un trabajo reciente en el que se analizan meta-análisis que miden el tamaño del efecto de esta metodología en comparación con otras más tradicionales, se pone de manifiesto que (Romero-Ariza, 2017):

- «[...] la cantidad de actividades enfocadas a promover la capacidad de pensar, razonar y elaborar argumentos basados en evidencias, es un predictor positivo de la comprensión de las ideas científicas por parte del alumnado sometido a aprendizaje por indagación (p. 291).
- [...] la indagación no guiada no facilita el aprendizaje mientras que cuando el proceso está debidamente asistido por el profesor sí se obtienen beneficios importantes (p. 292).
- [...] un tamaño del efecto moderado sobre el aprendizaje de conceptos y un mayor impacto sobre el desarrollo de destrezas de indagación [...] (p. 292).
- [...] el tamaño del efecto de la indagación depende de la clase de actividades de indagación que se han llevado a cabo en el aula, del nivel de guía o apoyo ofrecido por el docente y del tipo de resultado de aprendizaje medido (p. 292). »

Para arrojar luz a este asunto, en este artículo se realiza una revisión sistemática de los artículos publicados durante la última década en revistas españolas de DCE, siguiendo la metodología que se describe a continuación.

## **Método**

La revisión sistemática presentada en este trabajo ha sido realizada de acuerdo a:

- (1) La definición aportada por Higgins y Green (2008), en la que aquella se concibe como la revisión de una o varias preguntas claramente formuladas, usando unos métodos sistemáticos y explícitos para identificar y seleccionar las investigaciones referentes a ésta, y analizar los datos de aquellos estudios incluidos en la revisión.
- (2) La Declaración PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses), cuya finalidad es evaluar la calidad de revisiones sistemáticas y meta-análisis (Sotos-Prieto, Prieto, Manera, Baladia, Martínez-Rodríguez, Basulto, 2014).

La búsqueda se inició a principios de octubre de 2017, realizándose en las principales revistas españolas específicas del área de Didáctica de las Ciencias Experimentales: *Enseñanza de las Ciencias* (EC, en adelante), *Revista Eureka sobre Enseñanza y Divulgación de las Ciencias* (REEDC), *Revista Electrónica de Enseñanza de las Ciencias* (REEC), *Didáctica de las Ciencias Experimentales y Sociales* (DCES) y *Alambique*. Se revisaron los números publicados entre 2007 y 2017 (ambos inclusive), exceptuando los números extraordinarios de EC publicados en 2009 y 2013 por tratarse de las comunicaciones presentadas en los VIII y IX Congresos Internacionales sobre Investigación en DCE, respectivamente.

## Procedimiento de selección de artículos

La selección de los artículos fue realizada de forma independiente por cinco de los autores de este estudio, llevándose a cabo a través de las lecturas del título, resumen y palabras clave de cada trabajo. Se incluyeron estudios cualitativos y cuantitativos, además de experiencias y propuestas didácticas siempre que abordasen como uno de sus tópicos la enseñanza basada en indagación (o, en su defecto, enseñanza basada en investigación). En este sentido, los criterios de inclusión utilizados fueron:

- (1) Artículos publicados en revistas españolas especializadas en el área de Didáctica de las Ciencias Experimentales.
- (2) Artículos publicados entre 2007 y 2017, ambos inclusive.
- (3) Trabajos que aludan, bien en su título, resumen o palabras clave a la enseñanza basada en indagación.
- (4) Estudios que emiten implicaciones educativas referentes a la IBSE.

El proceso de selección de los trabajos, se realizó en dos fases. Primera, cada uno de los cinco autores encargados de dicha tarea revisó independientemente la revista que le fue adjudicada, detallando el número de artículos publicados por año y el de seleccionados de cada revista. Después, dos de los autores dedicados a esta labor, volvieron a revisar las cinco revistas, esta vez de forma conjunta, obteniendo un consenso del 91%. Los desacuerdos obtenidos en la segunda fase fueron resueltos consensuadamente por ambos investigadores.

## **Procedimiento de extracción de datos**

Para el análisis y la extracción de datos de los artículos se ha seguido el mismo procedimiento detallado en el apartado anterior. Así, los trabajos seleccionados se repartieron entre los mismos cinco autores encargados de la selección, realizando estos el proceso de forma independiente. Una vez finalizada la extracción de los datos en base a los siguientes ítems: (a) autor/es y año, (b) etapa educativa, (c) tipo de estudio, (d) diseño metodológico y (e) implicaciones educativas, estos fueron revisados de forma conjunta por los mismos dos autores encargados de la selección final, obteniendo un consenso del 84% respecto a la extracción inicial de datos. Los desacuerdos se resolvieron de forma consensuada por ambos investigadores.

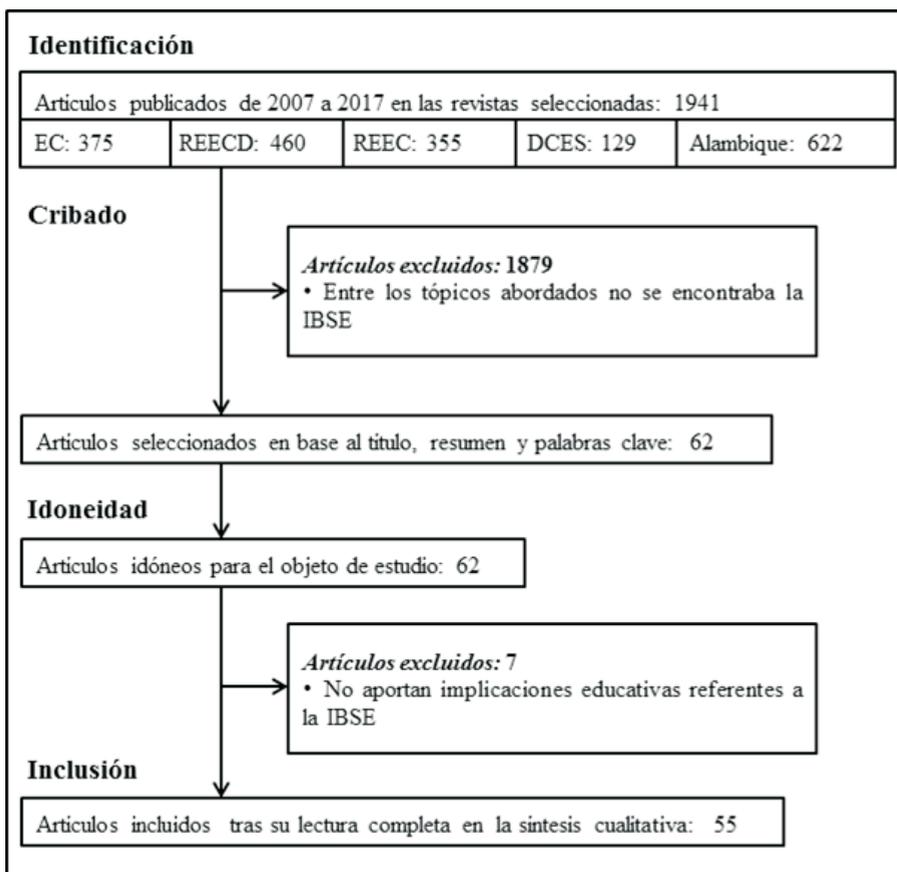
## **Resultados y Discusión**

Siguiendo las recomendaciones de la declaración PRISMA se ha elaborado el diagrama de flujo de la Figura II, en el que se muestra el proceso de selección e inclusión seguido en este estudio, según las acciones de identificación, cribado, idoneidad e inclusión.

Se revisaron 1941 artículos publicados entre los años 2007 y 2017 en las revistas españolas específicas de DCE. Tras la lectura del título, resumen y palabras clave de cada trabajo y la aplicación de los tres primeros criterios de inclusión fueron seleccionados 62 artículos, habiéndose excluido un total de 1879 que no abordaban (al menos explícitamente en las partes analizadas) la IBSE. El final de proceso de selección se realizó a través de la lectura completa de los 62 artículos

preseleccionados, de los cuales se excluyeron 7 por no cumplir con el cuarto criterio de inclusión, quedando finalmente una muestra de 55 trabajos lo que representa apenas un 3% de los trabajos revisados. Hecho esto, se procede a dar respuesta a los interrogantes de investigación.

FIGURA II. Diagrama de flujo de acuerdo a la declaración PRISMA.

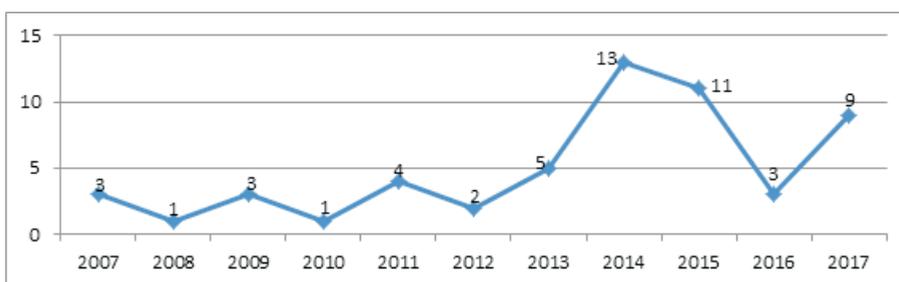


## Evolución de la producción de artículos sobre IBSE en España

¿Cuánto se investiga sobre IBSE? Si comparamos el número de artículos obtenido con el de otros trabajos similares, tales como los 138 trabajos

incluidos en la revisión de Minner et al. (2010) o los 164 seleccionados por Alfieri et al. (2011), podemos afirmar que la producción de trabajos publicados en revistas nacionales referentes a la IBSE resulta inferior a la internacional, puesto que ésta abarca un ámbito geográfico mucho más extenso al incluir solo artículos escritos en inglés y, normalmente, de países y revistas foráneas. La distribución de dicha producción se presenta en el Gráfico I.

GRÁFICO I. Evolución de la producción de artículos en España sobre IBSE.



Según puede observarse en el Gráfico I, la fluctuación de artículos referidos a la IBSE publicados en España es evidente. No obstante, se percibe cierto aumento en la producción, que queda refrendado por el hecho de que en los últimos cinco años se publica el 75% (N = 41) de los artículos seleccionados. Esto podría responder a las recomendaciones sobre el empleo de la IBSE emitidas tanto a nivel internacional (Ofsted, 2011; NRC, 2012) como nacional (Couso, Jiménez, López-Ruiz, Mans, Rodríguez, Rodríguez y Sanmartí, 2011).

Del total de trabajos seleccionados, el 84% han sido desarrollados en España, mientras que el 16% restante proviene de países de América Central y del Sur. Así, los autores de los nueve trabajos publicados en España pero realizados fuera de ella pertenecen a instituciones educativas de México (1), Brasil (5), Colombia (1), Argentina (1) y Bolivia (1). En este sentido, es importante señalar que no se ha encontrado ningún trabajo en el que colaboren instituciones educativas de diferentes países.

## Quién, en qué etapa educativa y cómo se investiga sobre IBSE

¿Quién investiga sobre IBSE? Se han identificado 82 autores entre los 55 artículos seleccionados, lo que representa un índice de colaboración de 1.49. A este hecho contribuye el número de estudios de un único autor, significando este el 33% (N = 18) de los artículos incluidos, resultado inferior al promedio obtenido para las Ciencias Sociales en España (2 autores por trabajo) y al establecido en la Didáctica de las Ciencias (1.81) (Anta y Pérez, 2007). Profundizando en el análisis sobre quién investiga en IBSE, se presentan las Tablas II y III.

TABLA II. Relación de autores y número de trabajos publicados sobre IBSE.

Clasificación	Nº de autores	Porcentaje
Autores con 1 trabajo	58	71%
Autores con 2 trabajos	15	18%
Autores con 3 trabajos	4	5%
Autores con 4 trabajos	4	5%
Autores con 5 o más trabajos	1	1%
Total	82	100%

Tal y como se observa en la Tabla II, el 71% de los autores tienen únicamente un trabajo, mientras que tan sólo el 11 % ostentan tres o más autorías. En este sentido, podríamos afirmar que no abundan grandes productores en esta línea de investigación (siempre basándonos en las publicaciones de las revistas españolas específicas de DCE) dado que el más productivo, J. Domènech Casal, ha publicado un total de 8 artículos en el lapso de tiempo estudiado. Además, el número de autores que han participado en la elaboración de 3 o más trabajos asciende a 9 (Tabla III). Esta situación, unida al escaso número de artículos sobre IBSE producidos en España, ya que tan sólo el 3% de los artículos publicados en las revistas analizadas entre 2007 y 2017 va dirigido a esta línea de investigación, podría apuntar a que la investigación educativa española sobre IBSE aún se encuentra en un estado incipiente.

**TABLA III.** Autores con mayor producción en España sobre IBSE

<b>Autor/a</b>	<b>Nº de artículos</b>
J. Domènech Casal	8
M. Martínez Chico R. López-Gay Lucio-Villegas M. R. Jiménez Liso B. Crujeiras Pérez	4
A. García Carmona C. Ferrés Gurt M. Romero Ariza I. M. Greca	3

¿En qué etapa educativa y cómo se investiga sobre IBSE? Las respuestas a esta cuestión las encontramos en la Tabla IV. En cuanto a la etapa educativa en la que se investiga, se han identificado trabajos desarrollados en: Educación Primaria, Educación Secundaria, Universidad y egresados (en referencia a la formación continua de los profesionales). No obstante, la atención prestada a cada etapa educativa está lejos de ser equitativa, pues la Educación Secundaria abarca casi el 50% de los estudios incluidos en la revisión, repartiéndose el resto entre las de Educación Primaria (12%), Universidad (29%) y egresados (4%). También se encuentran trabajos que no especifican la etapa educativa en la que se desarrolló el estudio o, en su caso, a la que va dirigida en el caso de las propuestas didácticas (2%), y otros en los que no procede hablar de etapa educativa por ser estudios de revisión de la literatura (7%). En consonancia, podemos hacer referencia a las conclusiones extraídas por Pro y Rodríguez (2010), dado que una vez más la investigación en DCE vuelve a centrarse en la etapa de Educación Secundaria, existiendo un número marginal o inapreciable de estudios en las etapas de Educación Infantil y Primaria, lo cual podría fundamentarse en la trayectoria profesional de los autores y autoras de los estudios seleccionados, pues muchos de ellos han ejercido o ejercen como profesores de Secundaria. Este hecho resulta contrario a las recomendaciones que nos llegan desde otras líneas de investigación, en las que se atribuyen múltiples beneficios al comenzar la educación científica en los primeros años de escolarización (Eshach y Fried, 2005).

De los resultados obtenidos puede afirmarse que la presencia de la metodología IBSE resulta casi inexistente en las aulas de ciencias en España

(Cortés et al., 2012). Seguramente, este hecho se deba principalmente a que los maestros tienen inseguridades sobre cómo implementar en el aula los procesos y métodos que utilizan habitualmente los científicos y/o mantienen, en muchas ocasiones, concepciones inapropiadas sobre lo que es la IBSE, tal como identificar la indagación como desviarse de la planificación curricular (Keys y Kennedy, 1999).

**TABLA IV.** En qué etapas educativas y cómo se investiga sobre IBSE

Ítem	VARIABLES	N	Porcentaje
Etapa educativa	Educación Primaria	7	12%
	Educación Secundaria	26	46%
	Universidad	16	29%
	Egresados (formación continua)	2	4%
	No específica	1	2%
	No procede	4	7%
Tipo de estudio	Trabajo empírico	23	42%
	Trabajo teórico	32	58%
Diseño	Experimental	6	11%
	Ex post-facto	3	5%
	Estudio de caso	12	21%
	Etnográfico	1	2%
	Investigación-acción	2	4%
	Revisión de la literatura	4	7%
	Experiencia didáctica	23	41%
	Propuesta didáctica	5	9%

Respecto al cómo se investiga sobre IBSE, nos hemos centrado en dos dimensiones: el tipo de estudio y el diseño del mismo. En cuanto al tipo de estudio, se observa que casi el 60% (N = 32) son trabajos teóricos en los que se describen experiencias o propuestas didácticas, o se presenta un aporte teórico a través de revisiones narrativas de la literatura. Consecuentemente, el número de estudios empíricos seleccionados supera escasamente la veintena. Respecto al diseño, el más repetido se corresponde con el de la descripción de una experiencia didáctica (N

= 23), mientras que el diseño empírico más utilizado es el estudio de caso (N = 12). Además, el número de estudios con diseño experimental es muy reducido; su presencia apenas supera el 10% de los artículos seleccionados.

A la luz de estos resultados, hemos de prestar atención a la escasez de estudios empíricos referidos a la metodología IBSE y publicados en las revistas españolas de DCE, dado que incide significativamente en la naturaleza de los resultados e implicaciones educativas obtenidas, mayormente resultantes de intuiciones, sensaciones u opiniones carentes del rigor y la sistematicidad necesarios para una posible réplica y generalización de los mismos. En este sentido, conviene apuntar que sería necesario implementar la metodología IBSE en las aulas, pero otorgándoles rigor metodológico, sistematicidad y reflexión, para de ese modo incrementar el número de estudios empíricos en esta línea de investigación. Esto podría hacerse bien a través de un diseño experimental basado en la evidencia y que nos permita cuantificar efectos (Bisquerra, 2014), o bien a través de diseños cualitativos como son el estudio de caso o la investigación-acción, no menos rigurosos y con la cualidad de mejorar la calidad de la práctica docente (en la enseñanza de las ciencias, en este caso) a partir de la misma (Elliot, 1993).

### **Qué implicaciones educativas se evidencian sobre IBSE**

Las implicaciones educativas que reflejan cada uno de los artículos seleccionados se encuentran sintetizadas en la Tabla V, en la que cada trabajo aparece referenciado a través del número asignado al mismo en el *Anexo*. Así, se han usado seis criterios para clasificarlas, referidos a: (1) competencias, (2) emociones, motivación y actitudes, (3) adquisición y transferencia del conocimiento, (4) naturaleza e imagen de la ciencia, (5) profesorado y (6) dificultades del alumnado.

**TABLA V.** Síntesis de aquellas implicaciones educativas ligadas a la IBSE.

Referidas a	N*	Implicaciones educativas
Competencias	3 7 3 8	<ul style="list-style-type: none"> <li>- Fomenta la autonomía del alumnado (6, 7, 23).</li> <li>- Incentiva el trabajo cooperativo entre el alumnado (1, 6, 18, 22, 23, 24, 43).</li> <li>- Desarrolla la competencia lingüística (16, 23, 43).</li> <li>- Favorece el desarrollo de habilidades relacionadas con la metodología científica (12, 13, 16, 18, 22, 33, 45, 48).</li> </ul>
Emociones, motivación y actitudes	4 15  2 1 1	<ul style="list-style-type: none"> <li>- Genera actitudes positivas hacia la ciencia (6, 24, 33, 44).</li> <li>- Eleva la motivación y la satisfacción hacia el aprendizaje de las ciencias (6, 11, 14, 16, 23, 30, 34, 35, 36, 37, 43, 46, 48, 52, 55).</li> <li>- El alumnado valora positivamente el empleo de esta metodología (32, 45).</li> <li>- Incrementa la autoestima del alumnado hacia el aprendizaje de las ciencias (24).</li> <li>- No incide en la motivación y estado emocional del alumnado (42).</li> </ul>
Adquisición y transferencia del conocimiento	11 5 9  5 1 1 4  2  6  1	<ul style="list-style-type: none"> <li>- Mejora la adquisición de contenidos (7, 12, 14, 15, 16, 23, 30, 38, 41, 42, 46).</li> <li>- Ayuda a alcanzar un aprendizaje significativo (10, 14, 15, 50, 52).</li> <li>- Fomenta la reflexión y el pensamiento crítico durante el proceso de enseñanza-aprendizaje (6, 10, 11, 17, 21, 22, 24, 46, 55).</li> <li>- Propicia la participación activa en el alumnado (24, 33, 34, 49, 55).</li> <li>- Aumenta el diálogo entre el docente y el alumnado (5).</li> <li>- Tiene en cuenta los intereses del alumnado y sus ideas previas (11).</li> <li>- Favorece la identificación y modificación de ideas erróneas por aquellas adecuadas al conocimiento científico (15, 24, 47, 50).</li> <li>- Ayuda a manifestar los diferentes problemas socio-científicos actuales y se adapta a las necesidades del alumnado, donde reina la incertidumbre generada por cambios extremadamente acelerados (25, 26).</li> <li>- Permite al alumnado tomar conciencia de la utilidad del conocimiento científico en su vida cotidiana (14, 28, 29, 43, 44, 46).</li> <li>- Facilita la transferencia de ideas y conocimientos de un contexto a otro (15).</li> </ul>
Naturaleza e imagen de la ciencia	4  3	<ul style="list-style-type: none"> <li>- Contribuye a presentar de forma más cercana la producción, desarrollo e impacto del conocimiento científico (53, 4, 5, 35).</li> <li>- Permite modificar concepciones negativas sobre la imagen de la ciencia y otorga valor a la labor científica (14, 29, 31).</li> </ul>
Profesorado	9  2 2 1  3 2  1  1	<ul style="list-style-type: none"> <li>- Formar a los futuros docentes en IBSE podría incidir en la innovación educativa futura, mejorando su conocimiento sobre dicha estrategia de enseñanza y generando más seguridad y mejor desempeño en el profesorado (9, 27, 31, 38, 39, 40, 41, 48, 49).</li> <li>- Requiere más dedicación por parte del docente que la enseñanza tradicional (33, 36).</li> <li>- El rol de guía del profesorado durante el proceso de indagación es esencial (19, 48).</li> <li>- El profesorado universitario considera prioritario emplear la IBSE en la formación inicial de maestros (39).</li> <li>- El profesorado en formación inicial valora positivamente a esta estrategia de enseñanza (1, 3, 51).</li> <li>- El profesorado en formación inicial tiene dificultades a la hora de diseñar secuencias didácticas basadas en indagación (3, 51).</li> <li>- Los maestros en formación poseen poco conocimiento sobre cómo se genera, valida y evoluciona el conocimiento científico; desconocen las técnicas y procedimientos básicos mediante los que se desarrolla el conocimiento científico, afectando todo ello a la imagen que pudieran transmitir (54).</li> <li>- El NPTAI es un instrumento de utilidad para el profesorado, pues se trata de un instrumento de evaluación formativa que permite identificar las dificultades del alumnado durante el proceso de indagación (20).</li> </ul>
Dificultades del alumnado	3  1 1  2  1	<ul style="list-style-type: none"> <li>- No posee un modelo claro de la actividad científica para poder ejecutar correctamente la dinámica de indagación (6, 8, 28).</li> <li>- Tiene dificultades en la formulación de preguntas investigables (21).</li> <li>- Presentan dificultades en la extracción de resultados, la interpretación de los mismos y la emisión de hipótesis (2).</li> <li>- Tanto la expresión oral y escrita como la identificación de las situaciones problemáticas, generalmente, son debilidades del alumnado (6, 36).</li> <li>- Afrontar actividades de indagación es complejo cuando tienen pocas vivencias relacionadas con este tipo de enfoque didáctico (19).</li> </ul>

\*Número de artículos que obtienen la implicación didáctica.

La contribución que realiza la metodología IBSE al desarrollo de competencias (autonomía e iniciativa, trabajo en equipo, competencia lingüística y aquellas referentes a la investigación) observada en muchos de los trabajos seleccionados en esta revisión es ratificada por Bevins y Price (2016), quienes afirman que una de las fortalezas de este enfoque de enseñanza es la promoción y el desarrollo de habilidades y capacidades ligadas a las ciencias. Además, estos mismos autores aportan como beneficios de la IBSE: la mejora de la autoestima; la comprensión de los contenidos y la motivación; y la incidencia en la aplicación de los nuevos aprendizajes en los contextos cada vez más complejos que les rodean.

En cuanto a la actitud y emociones hacia la Ciencia, se han encontrado resultados positivos y negativos, aunque mayoritariamente las aportaciones recogidas evidencian la eficacia de la IBSE para mejorar estas variables (Areepattamannil 2012; McConney et al. 2014). En este mismo sentido, las implicaciones educativas recogidas apuntan a una mejora de la adquisición de contenidos por parte del alumnado y al favorecimiento del aprendizaje significativo, actuando como responsables de ello la reflexión, la evaluación de las ideas alternativas, la argumentación (Pedaste et al. 2015) y la participación activa del alumnado (Bevins y Price, 2016); todas estas acciones se promocionan habitualmente por las propuestas y experiencias educativas analizadas.

Los efectos generados por la metodología IBSE descritos anteriormente, todos positivos, propician un clima de aula distendido, ameno y, consecuentemente, idóneo para el aprendizaje de las ciencias (Abril, Romero-Ariza, Quesada y García, 2014). A su vez, este grato beneficio es clave en la promoción de actitudes positivas hacia la escuela y el aprendizaje (Krüger, Formichella y Lekuona, 2015). Así, dado el declive de la actitud hacia la Ciencia que sufre el alumnado actual (Vázquez y Manassero, 2008), podríamos considerar a este enfoque de enseñanza como un medio adecuado para revertirlo.

Sin embargo, existen varios aspectos a tener en cuenta en la puesta en práctica de este enfoque de enseñanza según podemos observar en aquellas implicaciones educativas referentes al profesorado (Tabla V). Así, el rol que adquiere el docente durante el proceso de enseñanza-aprendizaje, su experiencia previa y conocimiento sobre IBSE y su implicación son factores determinantes en la eficacia de esta estrategia de enseñanza, convirtiendo al docente en un profesional eficaz (Reoyo, Carbonero y Martín, 2017). Tan es así, que algunos autores han cuestionado la metodología IBSE al alegar que:

- (1) El alumnado no puede por sí solo aprender por indagación (Kirschner, Sweller y Clark, 2006), pues existen evidencias que a través de una indagación guiada se obtienen mejores resultados en cuanto al aprendizaje se refiere (Minner et al. 2010; Furtak et al. 2012; Lazonder et al. 2016).
- (2) Falta preparación entre el profesorado (Forbes y David, 2010), dado el fragmentado, superficial y poco sólido conocimiento científico del que disponen (Murphy et al., 2007).
- (3) Las dificultades de los maestros en formación para diseñar y ejecutar secuencias basadas en indagación (Lucero et al., 2013; Crujeiras y Puig, 2014), línea sobre la que ya existen algunas propuestas de formación para intentar superar tales deficiencias (Martínez-Chico, Jiménez-Liso y López-Gay, 2015; Vílchez y Bravo, 2015; Greca, Meneses y Diez, 2017).

Finalmente, al hacer balance de las implicaciones educativas más y menos alentadoras, podemos afirmar que la enseñanza basada en indagación es una forma eficaz de enseñar y aprender ciencias, destacando como implicaciones más relevantes el incremento de la motivación/satisfacción por aprender ciencias y la adquisición de contenidos por parte del alumnado (N = 15 y N = 11 respectivamente). En este hecho coinciden Lederman, Lederman y Antink (2013), siempre que el diseño de las actividades sea adecuado, lo que pasa ineludiblemente por una adecuada formación del profesorado.

## Conclusiones

El presente estudio de revisión aporta, por una parte, una perspectiva del desarrollo de la investigación sobre la metodología IBSE en el área de DCE, centrada en las publicaciones de las revistas españolas; por otra, se recogen las principales implicaciones educativas referidas a dicha metodología. Así, podemos concluir lo siguiente:

Primero. La producción científica sobre la metodología IBSE en las revistas españolas específicas de DCE es escasa, a pesar de que existen autores de otros países que han difundido algunos trabajos en el nuestro. Esto contribuye a que resulte complejo resaltar algún investigador o grupo de investigación con una producción reseñable en esta línea de

investigación. Además, hay que destacar el exiguo grado de colaboración en la realización de los artículos. No obstante, existe una evolución favorable en cuanto al volumen de artículos que abordan dicha temática en los últimos cinco años. Así, podríamos afirmar que la investigación en España sobre la metodología IBSE es incipiente.

Segundo. La atención dedicada a las diferentes etapas educativas es muy desigual, siendo la etapa de Educación Secundaria y la universitaria a las que se dirige mayor número de estudios, lo que podría estar condicionado por la trayectoria profesional de los autores y autoras, que ejercen o han ejercido como profesores de Secundaria. Consecuentemente, la etapa de Educación Primaria recibe poca atención por parte de la investigación en DCE y, en este caso, no se destina ningún estudio a la Educación Infantil.

Tercero. Existen más estudios de naturaleza teórica que empírica. El diseño más empleado para aquellos trabajos teóricos ha sido la experiencia didáctica y para los estudios empíricos se ha optado, generalmente, por un diseño de estudio de caso. Por tanto, una vía para promover la extracción de conclusiones más rigurosas y generalizables sería incrementar el número de estudios experimentales (desde el punto de vista cuantitativo) o emplear en mayor medida la investigación-acción (desde el punto de vista cualitativo).

Cuarto, y último. Las implicaciones educativas referentes a la metodología IBSE son alentadoras en cuanto a los efectos generados en el alumnado y profesorado en formación, pues en ambos posibilita una sólida alfabetización científica apoyada en actitudes y emociones positivas, y en una imagen de la ciencia renovada. Sin embargo, este enfoque de enseñanza tiene ante sí algunos retos importantes:

- Referentes al profesorado:
  - Mejorar el conocimiento de los futuros maestros sobre naturaleza de la ciencia.
  - Paliar las dificultades de estos en el desempeño de la metodología IBSE, que giran en torno a la inseguridad generada por lo descrito anteriormente y la falta de experiencia en la puesta en práctica de dicha metodología.
- Referentes al alumnado:
  - Promover un aprendizaje de las ciencias basado en hacer ciencia, en aras de corregir actitudes deterioradas, sensaciones y emociones negativas y concepciones erróneas sobre las ciencias.

De las conclusiones obtenidas parece desprenderse que la IBSE podría ser una buena estrategia de enseñanza-aprendizaje de las ciencias, y también que queda mucho camino que recorrer. Son estudios como los presentados en este artículo los que pueden delimitar el camino a seguir. Dadas las limitaciones del mismo, pues se centra en investigaciones de ámbito nacional (España), la línea a seguir pasaría, en primer lugar, por ampliar el estudio al ámbito internacional, lo que se tiene previsto hacer en futuras revisiones.

## Agradecimientos

Este trabajo se enmarca dentro del proyecto «Diseño, evaluación y difusión de talleres basados en indagación para Educación Primaria (ref: PPJI\_B-01)» financiado por la Universidad de Granada. Además hemos de agradecer a la Junta de Andalucía y al Fondo Social Europeo por la financiación del contrato nº 6161 destinado a la contratación de joven personal investigador, al Ministerio de Educación, Cultura y Deporte por la concesión de la beca FPU15/04972 y al grupo de investigación HUM-613, por participar en su formación.

**Anexo:** [www.didacticacienciasugr.es/docs/misc/Anexo-Rev-Sis-Nacional.pdf](http://www.didacticacienciasugr.es/docs/misc/Anexo-Rev-Sis-Nacional.pdf)

## Referencias bibliográficas

- Abril, A. M., Romero-Ariza, M., Quesada, A., & García, F. J. (2014). Creencias del profesorado en ejercicio y en formación sobre el aprendizaje por investigación. *Revista Eureka sobre enseñanza y divulgación de las Ciencias*, 11(1), 22-33.
- Alfieri, L., Brooks, P.J., Aldrich N.J., & Tenenbaum H.R. (2011). Does discovery-based instruction enhance learning? *Journal of Educational Psychology*, 103, 1-18.
- Anderson, R.D. (1996). *Study of curriculum reform*. Washintong DC: U.S. Government Printing Office.

- Anta, C. & Pérez, J. M. (2007, noviembre, 21-23). *La producción científica sobre Didáctica de las Ciencias. IV Congreso Comunicación Social de la Ciencia*. Madrid: CSIC.
- Areepattamannil, S. (2012). Effects of inquiry-based science instruction on science achievement and interest in science: Evidence from Qatar. *The Journal of Educational Research*, 105(2), 134–146.
- Barrow, L. H. (2006). A brief history of inquiry: from Dewey to Standards. *Journal of Science Teacher Education*, 17, 265-278.
- Bevins, S. & Price Centre for Science Education, Sheffield Hallam University, Sheffield, UK, G. (2016). Reconceptualising inquiry in science education. *International Journal of Science Education*, 38(1), 17-29.
- Bisquerra, R. (Coord.) (2014). *Metodología de investigación educativa*. Madrid: La Muralla.
- Cañal de León, P. (1987). Un enfoque curricular basado en la investigación. *Investigación en la Escuela*, 1, 43-50.
- Cobern, W.W., Schuster, D., Adams, B., Applegate, B., Skjold, B., Undrieu, A., Loving, C.C., & Gobert, J.D. (2010). Experimental comparison of inquiry and direct instruction in science. *Research in Science & Technological Education*, 28(1), 81-96.
- Cortés, A.L., de la Gándara, M., Calvo, J.M., Martínez, M.B., Gil, M.J., Ibarra, J. & Arlegui, J. (2012). Expectativas, necesidades y oportunidades de los maestros en formación ante la enseñanza de las ciencias en la educación primaria. *Enseñanza de las Ciencias*, 30(3), 155-176.
- Couso D., Jiménez M. P., López-Ruiz J., Mans C., Rodríguez C., Rodríguez J.M., & Sanmartí, N. (2011). *Informe ENCIENDE: Enseñanza de las Ciencias en la Didáctica escolar para edades tempranas en España*. Madrid: Rubes Editorial.
- Couso, D. (2014). *De la moda de “aprender indagando” a la indagación para modelizar: una reflexión crítica*. Ponencia a los XXVI Encuentros de Didáctica de las Ciencias Experimentales, Huelva.
- Crawford, B.E. (2007). Learning to teach science as inquiry in the rough and tumble of practice. *Journal of Research in Science Teaching*, 44, 613–642.
- Crujeiras, B. & Puig, B. (2014). Trabajar la naturaleza de la ciencia en la formación inicial del profesorado planificando una investigación. *Educació Química*, 17, 55-61.

- Cuevas, P., Lee, O., Hart, J. & Deakort, R. (2005). Improving science inquiry with elementary students of diverse backgrounds. *Journal of Research in Science Teaching*, 42, 337-357.
- Elliot, J. (1993). *El cambio educativo desde la investigación-acción*. Madrid: Morata.
- Eshach, H. & Fried M. N. (2005). Should science be taught in early childhood? *Journal of Science Education and Technology*, 14(3), 315-336.
- Forbes, C. T. & Davis, E. A. (2010). Curriculum design for inquiry. Pre-service elementary teachers' mobilization and adaptation of science curriculum materials. *Journal of Research in Science Teaching*, 47, 820-839.
- Furtak, E.M., Seidel, T., Iverson, H., & Briggs, D.C. (2012). Experimental and Quasi-Experimental Studies of Inquiry-Based Science Teaching. *Review of Educational Research*, 82(3), 300-329.
- Gil-Pérez D. (1990). *Un modelo de resolución de problemas como investigación*. Madrid: Ministerio de Educación y Ciencia – Lábor.
- Gil-Pérez D. (1993). Contribución de la historia y la filosofía de las ciencias al desarrollo de un modelo de enseñanza/aprendizaje como investigación. *Enseñanza de las Ciencias*, 11(2), 197-212.
- Greca, I. M., Meneses, J. A., & Díez, M. (2017). La formación en ciencias de los estudiantes del grado en maestro de Educación Primaria. *Revista Electrónica de Enseñanza de Las Ciencias*, 16(2), 231-256.
- Hazelkorn, E. (2015). *Science Education for Responsible Citizenship*. Comisión Europea. [http://ec.europa.eu/research\\_swafs\\_pdf\\_pub\\_science\\_education\\_KI-NA-26-893-EN-N](http://ec.europa.eu/research_swafs_pdf_pub_science_education_KI-NA-26-893-EN-N)
- Higgins, J. P. T., & Green, S. (2008). *Cochrane handbook for systematic reviews of interventions*. West Sussex, England: Cochrane Collaboration & Wiley.
- Keys, C.W. & Kennedy, V. (1999). Understanding inquiry science teaching in context: A case study of an elementary teacher. *Journal of Science Teacher Education*, 10(4), 315-333.
- Krüger, N., Formichella, M. M., & Lekuona, A. (2015). Más allá de los logros cognitivos: la actitud hacia la escuela y sus determinantes en España según PISA 2009. *Revista de Educación*, 367, 10-35.
- Kirschner P.A., Sweller J., & Clark R.E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist*, 41, 75-86.

- Lazonder, A.W., & Harmsen, R. (2016). Meta-Analysis of Inquiry-Based Learning Effects of Guidance. *Review of Educational Research*, 20(10), 1-38.
- Linn, M. C., Davis, E.A. & Bell, P. (2004). *Internet environments for science education*. New Jersey: Lawrence Erlbaum Associates.
- Lucero, M., Valcke, M. & Schellens, T. (2013). Teachers' beliefs and self-reported use of inquiry in Science Education in Public Primary Schools. *International Journal of Science Education*, 35 (8), 1407-1423.
- Martínez-Aznar, M. (2017). La enseñanza-aprendizaje de las ciencias: «entre indagación y problemas». Comunicación personal.
- Martínez-Chico, M. (2013). *Formación inicial de maestros para la enseñanza de las ciencias. Diseño, implementación y evaluación de una propuesta de enseñanza*. Tesis doctoral. Universidad de Almería.
- Martínez-Chico, M., Jiménez Liso, M. R., & López-Gay, R. (2015). Efecto de un programa formativo para enseñar ciencias por indagación basada en modelos, en las concepciones didácticas de los futuros maestros. *Revista Eureka Sobre Enseñanza Y Divulgación de Las Ciencias*, 12(1), 149-166.
- McConney A., Oliver M.C., Woods-McConney A., Schibeci R. & Maor D. (2014) Inquiry, Engagement, and Literacy in Science: A Retrospective, Cross-National Analysis Using PISA 2006. *Science Education*, 98(6), 963-980.
- McDonald, S. & Butler Songer, N. (2008). Enacting classroom inquiry: Theorizing teachers' conceptions on science teaching. *Science Education*, 27, 45-60.
- Minner, D.D., Levy, A.J. & Century, J. (2010). Inquiry-Based Science Instruction—What Is It and Does It Matter? Results from a Research Synthesis Years 1984 to 2002. *Journal of Research in Science Teaching*, 47(4), 474-496.
- Murphy, C., Neil, P. & Beggs, J. (2007) Primary science teacher confidence revisited: ten years on. *Educational Research*, 49(4), 415-430.
- National Research Council. (2012). *A Framework for K-12 science education: Practices crosscutting concepts, and core ideas*. Washington, DC: National Academy Press.
- Ofsted (2011). *Successful science: An evaluation of science education in England 2007 – 2010*. Manchester, UK: Ofsted.
- Pedaste M., Mäeots M., Siiman L.A., De Jong T., Van Riesen S.A., Kamp E.T., & Tsourlidaki E. (2015). Phases of inquiry-based learning: Definitions and the inquiry cycle. *Educational research review*, 14, 47-61.

- Perales, F.J., & Ayerbe, J. (2016). El trabajo por proyectos y por resolución de problemas en Educación Ambiental: análisis y tendencias. Comunicación a los *XXVII Encuentros de Didáctica de las Ciencias Experimentales*. Badajoz (7-9 de Septiembre, 2016).
- Pro, A., & Rodríguez, (2010). Aprender competencias en una propuesta para la enseñanza de los circuitos eléctricos en Educación Primaria. *Enseñanza de las Ciencias*, 28(3), 385-406.
- Reoyo, N., Carbonero, M. A., & Martín, L. J. (2017). Características de eficacia docente desde las perspectivas del profesorado y futuro profesorado de secundaria. *Revista de Educación*, 376, 62-86.
- Rocard, M. (Chair), Csermely, P., Jorde, D. Lenzen, D., Walberg-Henriksson, H. & Hemmo, V. (2007). *Science Education now: A renewed Pedagogy for the future of Europe*. Comisión Europea. <https://sede.educacion.gob.es/publivera/PdfServlet?pdf=VP15136.pdf&area=E>
- Sotos-Prieto, M., Prieto, J., Manera, M., Baladia, E., Martínez-Rodríguez, R. & Basulto, J. (2014). Ítems de referencia para publicar Revisiones Sistemáticas y Metaanálisis: La Declaración PRISMA. *Revista Española de Nutrición Humana y Dietética*, 18(3), 172-181.
- Vázquez, A. & Manassero, M.A. (2008). El declive de las actitudes hacia la ciencia de los estudiantes: un indicador inquietante para la educación científica. *Revista Eureka Enseñanza y Divulgación de las Ciencias*, 5(3), 274-292.
- Vílchez, J. M., & Bravo, B. (2015). Percepción del profesorado de ciencias de educación primaria en formación acerca de las etapas y acciones necesarias para realizar una indagación escolar. *Enseñanza de Las Ciencias*, 33(1), 185–202.
- Windschitl, M. (2003). Inquiry projects in science teacher education: What can investigative experiences reveal about teacher thinking and eventual classroom practice? *Science Education*, 87(1), pp. 112-143.
- Yeomans, E. (2011). *Inquiry-based learning – what is its role in an inspiring science education?* En *Perspectives on Education: Inquiry-based learning* (p. 3). Wellcome Trust: Londres.

**Información de contacto:** David Aguilera Morales. Facultad de Ciencias de la Educación, Campus Universitario de Cartuja, C.P. 18071 (Granada), Granada. E-mail: david15@correo.ugr.es

# ¿Mejoran los proyectos de divulgación con experimentación la actitud hacia las clases de ciencias?!

## Do experimentation outreach programs improve the attitudes towards school science?

DOI: 10.4438/1988-592X-RE-2017-381-389

Raquel Fernández-César

Natalia Pinto-Solano

*Universidad de Castilla la Mancha*

Marta Muñoz-Hernández

*Universidad Rey Juan Carlos*

### Resumen

Existe un interés creciente por el dominio afectivo en la didáctica de las ciencias, y en concreto por la actitud hacia las clases de ciencias como posible acción contra el declive de vocaciones científicas que se viene observando en los últimos años. Por ello, este trabajo valora el efecto de un programa divulgativo basado en la experimentación sobre la actitud hacia las clases de ciencias. En dicha valoración se tienen en cuenta factores como la etapa educativa, el sexo y el entorno del centro (rural o urbano).

La metodología utilizada es cuantitativa, el estudio es transversal, y el instrumento para medir la actitud es un cuestionario breve que se administra como pretest y postest. La muestra está compuesta por 380 alumnos de los centros escolares no universitarios participantes en el proyecto. El análisis estadístico se realiza con el paquete informático SPSSv22.

Los resultados obtenidos al comparar el pretest y el postest muestran: una actitud más positiva hacia las ciencias en la escuela en el postest ( $t=5.46$ ;  $gl=299$ ;  $p<.01$ ), mayor en el alumnado de la etapa de educación primaria ( $t=-6.61$ ;  $gl=210$ ;

---

<sup>(1)</sup> Los autores agradecen la financiación de la Fundación Española de Ciencia y Tecnología, FECYT, en convocatorias de 2012 y 2014, bajo los proyectos FCT-12-4221 y FCT-14-8844.

$p < .00$ ), y en el alumnado de zonas rurales ( $p < .00$ ); no se aprecia influencia del sexo en el efecto del programa sobre la actitud ( $p = .38$  antes;  $p = .06$  después).

Se concluye que la actuación desarrollada tiene una influencia más positiva en los estudiantes de la etapa de educación primaria que en los de educación secundaria. Se propone que las acciones que pretendan actuar sobre los futuros itinerarios científicos a seguir por los estudiantes deberían iniciarse en la etapa de primaria, donde se obtiene un efecto más positivo, y prolongarse e intensificarse en educación secundaria. Es necesaria una investigación más profunda para confirmar (o no) si en las zonas rurales el alumnado presenta una actitud hacia las ciencias más positiva de forma generalizada.

**Palabras clave:** Actitud hacia las clases de ciencias, Educación Primaria, Educación Secundaria, Programa de divulgación

### **Abstract**

There is an increasing interest in the affective domain in science education, and particularly in the attitudes towards school science, as a possibility to counteract the present decline of scientific vocations observed in the last few years. Thus, this study aims to evaluate the effect of an experimental outreach program on the attitude towards school science. The possible influence of factors such as educational stage, sex and geographical context (rural or urban) is taken into account.

This is a quantitative, cross-sectional study with a brief questionnaire being administered before and after the intervention to measure student attitude. The sample consists of 380 students from the pre-university schools participating in the project. Statistical analysis was carried out using the SPSSv22 computer software.

The results obtained after comparing the pretest and posttest show a more positive attitude towards school science in the posttest ( $t = 5.46$ ;  $df = 299$ ;  $p < .01$ ). This increase is more prevalent among elementary school students ( $t = 6.61$ ;  $df = 210$ ;  $p < .00$ ), and among students from rural areas ( $p < .00$ ). No significant difference based on sex is found ( $p = .38$  pretest;  $p = .06$  posttest).

The conclusion of our evaluation is that this outreach program has a more positive influence on elementary students than on those in secondary education. Therefore, interventions regarding the future selection of science related subjects should take place in the early stages of education, such as the elementary level, where a more positive increase is obtained, and to keep the actions in a more intense and long-term format. With regard to the increase in positive attitudes among students living in rural environments found in this study, further research is needed to confirm whether or not these results can be generalized.

**Key words:** Attitude towards school science, Elementary Education, Secondary Education, Outreach program

## Introducción

El estudio sobre las actitudes se desarrolla inicialmente en el campo de la psicología social (Allport, 1935, citado en di Martino & Zan, 2010) con el objetivo de prever las elecciones individuales relacionadas, entre otros, con la intención de voto y la compra de artículos. La actitud se manifiesta a través de componentes conductuales, cognitivos y afectivos (Heckman, Stixrud & Urzua, 2006) que pueden ser positivos o negativos en el sentido de expresar atracción o aversión (Eiser, 1989), además, son susceptibles de ser modificados ya que la actitud no es un rasgo estable (Briñol et al, 2004).

En el ámbito de la educación la actitud es un constructo que a lo largo del tiempo ha ido respondiendo, según el paradigma existente, a tres modelos conceptuales: el unidimensional, el de dos componentes (afectiva y cognitiva) y el tripartito (comportamientos, emociones y cogniciones). Las evidencias soportadas por el ámbito de ecuaciones estructurales han confirmado que el de dos componentes es el que presenta un mejor ajuste de modelo, significando, por lo tanto, que la actitud está constituida por variables afectivas y cognitivas que se generan a partir de la experiencia personal de los sujetos (Vázquez, 2013). El trabajo que se presenta se enmarca en dicho modelo bidimensional donde se considera que la actitud hacia las clases de ciencias engloba cogniciones y afectos.

En el campo de la didáctica de las ciencias no ha existido una tradición de desarrollo de modelos de las actitudes, sino más bien de realización de medidas de las mismas con diversos instrumentos. Ya en ese siglo se han realizado esfuerzos para dotar a este campo de esa fundamentación teórica (Alsop, 2005; Saleh & Khine, 2011). También en este siglo se han llevado a cabo revisiones de los estudios realizados con diversos instrumentos en el tema de actitudes hacia las ciencias (Osborne, Simon, & Collins, 2003; Mellado et al., 2013; Potvin & Hasni, 2014). En dichas revisiones se reconoce, dentro de la taxonomía de la actitud hacia las ciencias, la categoría asignada a la ciencia en la escuela o las clases de ciencias. Por lo tanto, los autores de esos informes consideran que una parte de la actitud que los alumnos tienen sobre la ciencia se forma en sus clases de ciencias. Desde el punto de vista de Vázquez- Alonso y Manassero- Mas, (2011) es necesaria la inclusión de la actitud hacia las clases de ciencias en cualquier modelo teórico que pretenda explicar y/o mejorar los procesos de enseñanza- aprendizaje de las ciencias en

la escuela. En esta línea de imbricación de la actitud y la enseñanza-aprendizaje de las ciencias los resultados son contradictorios al respecto de la experimentación como parte de dicha instrucción. Mientras Vázquez- Bernal y Jiménez- Perez (2016) señalan la existencia de una asociación entre la utilización de metodología experimental en el aula y la actitud del alumnado hacia las clases de ciencias, los resultados de (Fernández Cézar & Solano Pinto, 2017) no reportan la existencia de dicha asociación.

La mayor parte de los estudios sobre la actitud hacia las ciencias emplean cuestionarios como herramienta para su valoración (Rubba, & Harkness, 1993; Manassero-Mas, Vázquez-Alonso, & Acevedo-Díaz, 2001; Manassero, Vázquez, & Acevedo, 2003; Siegel, & Ranney, 2003; Chen, 2006; Vázquez Alonso, Acevedo Díaz, Manassero Mas, & Acevedo Romero, 2006; Vázquez Marbà-Tallada, & Márquez Bargalló, 2010; de Pro Bueno & Pérez Manzano, 2014; Kennedy, Quinn & Taylor, 2016). La relación entre la información obtenida y el tiempo empleado se optimiza con estos instrumentos, como se reporta en los trabajos de Vázquez, Manassero y Acevedo (2006), y más recientemente en el Kennedy, Quinn y Taylor (2016). Existen cuestionarios realizados para medir la actitud hacia las ciencias con orientación hacia la ciencia, tecnología y sociedad, CTS, como el COCTS (Bennássar, Vázquez, Manassero & García-Carmona, 2010) compuesto por preguntas tomadas del cuestionario Views on Science, Technology and Society, VOST, (Aikenhead, & Ryan, 1992). También existen estudios de amplio espectro como el estudio de la relevancia de la educación científica, The Relevance of Science Education, ROSE (Schreiner & Sjøberg, 2004), que se ha realizado a escala europea con alumnado de 15 a 17 años. Los datos de España para este último estudio están publicados en diversos artículos (Alonso, & Mas, 2007; Alonso, & Mas, 2009; ROSE, A. C. D. P., 2005). Pero todos estos estudios emplean cuestionarios muy largos, con preguntas de formulación relativamente compleja, y de difícil comprensión, incluso para los estudiantes a los que van dirigidos. Además, con un número excesivamente grande de ítems, 657 en el caso del VOST, que impide que sean utilizados autónomamente por el alumnado joven, como el que nos interesa en este trabajo. El empleo de muchos ítems puede provocar confusión y abandono por parte de los estudiantes evaluados, la presencia de un gran número de respuestas incorrectas debido al cansancio, así como el empleo de grandes cantidades de tiempo para recabar la información a los investigadores

(Abad, Olea, Ponsoda & García, 2011). Se trata, por lo tanto, de utilizar un número de ítems tal que nos permita encontrar una información adecuada con una relación número de ítems/tiempo asumible.

En línea con este objetivo, se han encontrado algunos estudios actuales orientados al análisis de las actitudes hacia las clases de ciencias con escolares de edades inferiores a 12 años (Kennedy, Quinn & Taylor, 2016; Fernández Cézar & Solano Pinto, 2017) donde se opta por reducir el número de preguntas y simplificar la formulación de las mismas para que puedan ser completados por los alumnos de manera rápida y autónoma. En estos cuestionarios se mantiene lo que según, di Martino y Zan (2010), son los tres aspectos fundamentales de la actitud: *me gustan / no me gustan* (interés-afectivo), porque *puedo/no puedo hacer* (percepción de la propia competencia-cognitivo), porque *las ciencias son/no son* (visión de las ciencias-cognitivo).

Se considera la actitud como un conjunto de cogniciones y emociones modificables mediante acciones, y por tanto, la actitud hacia las clases de ciencias podría ser susceptible de mejorarse. En este trabajo se contempla una acción consistente en un proyecto de fomento de vocaciones científicas, accesible en [www.nanomadrid.es](http://www.nanomadrid.es), y consistente en el acercamiento a los escolares de centros docentes no universitarios de experimentos científicos llevados a cabo por investigadores de distintas universidades y centros de investigación españoles, junto con los propios estudiantes. Los experimentos solicitados por los centros docentes tocan temáticas curriculares del nivel educativo para el que los solicitan, como las reacciones químicas, cambios de estado, propiedades de materiales (magnetismo, electricidad...), superconductividad, etc. Según la etapa educativa y el momento evolutivo de los participantes, se adapta el contenido y la metodología, siempre conservando la esencia del aprendizaje por descubrimiento involucrando a los estudiantes en la implementación del método científico. El planteamiento incluye experimentación por parte de los escolares, y propicia su reflexión sobre el tema que se trata en cada caso, pues no hay aprendizaje sin elaboración de la propia información mediante la reflexión.

Existen estudios previos en los que se valora la influencia de distintas acciones sobre la actitud hacia las clases de ciencias de alumnado preuniversitario (Barmby, Kind, & Jones, 2008; Cardona et al., 2012). Se encuentran estudios más recientes sobre medidas de la actitud en alumnado de edades inferiores a 12 años, pero las acciones que analizan se

refieren a cambios en la instrucción provocados por el uso de Webquests (Unal, & Altiparmak, 2013), a programas relacionados con la percepción social de la ciencia (Hong, Lin, Wang, Chen, & Yang, 2013), o a la adaptación completa del currículo de ciencias al empleo de determinado material tecnológico como los legos (Bethke Wendell, & Rogers, 2013). Ninguno de estos trabajos contempla la inclusión de los experimentos como acción principal sobre la actitud hacia las clases de ciencias en la escuela. Aunque en algún trabajo anterior se reporta la influencia positiva sobre la motivación hacia las ciencias y las matemáticas de la inclusión de la experimentación en la instrucción educativa (Weinberg, Basile, & Albright, 2011), que sería lo deseable desde el punto de vista didáctico, el trabajo que aquí se presenta difiere del anterior en que las acciones son llevadas a cabo con los estudiantes por investigadores, personas distintas al docente habitual.

Por todo ello, en el presente trabajo se plantea como objetivo general la valoración de si intervenciones basadas en la experimentación con la participación de los alumnos de Educación Primaria (EP) y Secundaria (ES) mejoran su actitud hacia las clases de ciencias. Además, como objetivos específicos, se plantea el estudio de las posibles diferencias sobre esta actitud relacionadas con algunos factores: sexo, etapa educativa (EP y ES), y entorno donde se encuentra el centro educativo (rural y urbano).

Teniendo en cuenta la bibliografía revisada se espera encontrar los siguientes resultados: La participación del alumnado en los experimentos mejora la actitud hacia la clase de ciencias; dicha mejora en la actitud es mayor en la etapa educativa de EP; las mayores puntuaciones las obtienen los jóvenes de sexo masculino; y el alumnado del entorno rural presenta mayores cambios en su actitud.

## Método

La metodología empleada es cuantitativa: Se mide la actitud hacia las clases de ciencias en la escuela mediante un cuestionario. Se trabaja con alumnado del último año de educación primaria, 3º de ESO y 1º de bachillerato, por lo que se trata de un estudio transversal que nos proporciona una instantánea de la situación existente en los cursos 2014-2015 y 2015-2016.

## Muestra

La muestra es de cercanía, y está compuesta por estudiantes de los centros educativos que han solicitado participar en el proyecto de *Conocer la ciencia hoy abre las puertas del mañana* (Fernández César & Solano Pinto, 2017), que se viene desarrollando desde 2012 por un grupo de científicos de varias universidades y centros de investigación de Madrid y Castilla La Mancha. De los 500 estudiantes con los que se ha trabajado, 380 cuestionarios se consideran válidos pues se han completado adecuadamente, y constituyen la muestra total. De ellos, 154 son chicas y 146 chicos; 199 viven en zonas urbanas y 101 en rurales (pueblos de menos de 5000 habitantes); 227 son de último curso de educación primaria, 68 de 3º curso ESO, y 85 de 1º curso de bachillerato de Ciencias de la salud-tecnológico. Una descripción más detallada de la muestra se incluye en la tabla I.

TABLA I. Descripción de la muestra por sexos y niveles educativos

Sexo		Zona urbana	Zona rural	Total
Chica	Primaria	77	31	108
	ESO	34	10	44
	Bachillerato	18	22	40
	Total	139	53	92
Chico	Primaria	62	57	119
	ESO	18	6	24
	Bachillerato	19	26	45
	Total	105	83	188

## Instrumento

El cuestionario lo componen 8 ítems tomados del dominio “My Science Class” del cuestionario ROSE (Schreiner & Sjøberg, 2004), empleado en un estudio previo con alumnado de primaria (10-12 años) (Fernández César & Solano Pinto, 2017). Las respuestas son tipo escala Likert de 1 a 5, donde 1 es totalmente en desacuerdo (TD), 3 es neutro y 5 es

totalmente de acuerdo (TA). Están formulados de manera que se evite la aquiescencia en las respuestas. Para cuantificar la actitud total, se hace coincidir en todos ellos el 5 con el máximo y el 1 con el mínimo. En la tabla II se muestran los ítems. La consistencia interna del conjunto de ítems se mide mediante alfa de Cronbach, y está publicada en un trabajo previo con una muestra con menos participantes (Fernández Cézar & Solano Pinto, 2017). No obstante, se ha vuelto a valorar con esta muestra obteniendo un valor de .810, dentro del rango de valores aceptables para ciencias sociales.

---

TABLA II. Cuestionario empleado

---

- |  |
|--|
| <ol style="list-style-type: none"><li>1. Las Ciencias naturales <b>son</b> una asignatura difícil</li><li>2. La clase de ciencias <b>es</b> interesante</li><li>3. La clase de Ciencias ha abierto mis ojos a trabajos interesantes que <b>podría hacer cuando sea adulto</b></li><li>4. <b>Me gustan</b> las Ciencias naturales más que otras asignaturas</li><li>5. Creo que <b>todo el mundo debería estudiar</b> Ciencias en la Escuela</li><li>6. Las cosas que aprendo en la clase de Ciencias <b>me será de ayuda</b> en mi vida diaria</li><li>7. La clase de ciencias <b>me ha mostrado</b> la importancia de la ciencia en nuestro estilo de vida</li><li>8. <b>Me gustaría tener</b> cuanta más ciencia mejor en la Escuela</li></ol> |
|--|

## Procedimiento

Los cuestionarios se envían a los colegios previa solicitud de participación de los mismos en el proyecto. Se administran antes y después de la intervención con los estudiantes, previa obtención del consentimiento informado. El pretest es realizado por los investigadores que llevan a cabo el programa antes de realizar el experimento. Una semana después de la realización del mismo, el tutor, que ha sido instruido para realizar la evaluación, lo vuelve a administrar a los escolares y envía las respuestas, que constituyen el postest.

La intervención se realiza en horario escolar, y consiste en una sesión de entre 1h y 2h de duración. Durante la misma los estudiantes trabajan en parejas o en grupos de 4 como máximo, realizando el experimento que el colegio haya solicitado de entre los que se ofrecen en el programa.

No se les proporciona ningún guion cerrado, sino que se sigue el método científico: los investigadores, y a veces los estudiantes, plantean preguntas, y los estudiantes con sus materiales y conocimientos emiten hipótesis o propuestas de respuestas que validan ellos mismos. Como todos los grupos trabajan sobre el mismo problema o cuestión, finalmente se llega a una conclusión que se comparte con toda la clase.

El análisis de los datos se lleva a cabo con el paquete informático SPSSv22. Se suman los 8 ítems para obtener la actitud total, llamada AT. Se justifica el uso del valor total como medida de la actitud mediante los valores de la correlación ítem-total, y mediante análisis factorial exploratorio con matriz de covarianzas obtenida con método varimax con normalización de Kaiser.

Para diferenciar la actitud total antes y después de la participación en el proyecto, se le añade la letra A al momento anterior a la intervención, y D para indicar después de la misma, quedando ATA y ATD, respectivamente. Estas son variables cuantitativas, y los factores que se analizan son la etapa (primaria, secundaria obligatoria o bachillerato), el sexo y la zona donde está ubicado el centro escolar (rural o urbana). Con el análisis de normalidad de las variables cuantitativas se determina el tipo de contraste a realizar. La comparación de la ATA y ATD se realizará mediante contraste de medias para muestras relacionadas. En todos los casos, el intervalo de confianza de los análisis es del 95%, si no se indica otra cosa.

## Resultados

Se analiza la actitud total antes y después de la participación en el proyecto. Los valores de  $\rho$  de Spearman para la correlación ítem-total se muestran en la tabla III, a.

**TABLA III:** a) Correlación ítem-total para la actitud antes y después de la intervención

	<b>ATA</b>	<b>ATD</b>
ITEM1	.53	.58
ITEM2	.66	.56
ITEM 3	.60	.66
ITEM 4	.60	.60
ITEM5	.52	.53
ITEM6	.64	.66
ITEM7	.57	.65
ITEM8	.56	.43

b) Cargas factoriales de cada ítem al modelo de un factor (F1)

ITEM1	.42
ITEM2	.57
ITEM 3	.51
ITEM 4	.46
ITEM5	.64
ITEM6	.53
ITEM7	.46
ITEM8	.38

El análisis factorial exploratorio (K.M.O = .743;  $p < .000$  para Esfericidad de Bartlett) devuelve un modelo de un factor que explica el 34.43% de la varianza, con cargas factoriales extraídas con el método de máxima verosimilitud para la matriz de covarianzas obtenida con método varimax con normalización de Kaiser, cuyos valores se muestran en la tabla III b.

Tanto la correlación ítem-total como el análisis factorial exploratorio justifican el empleo de la actitud total, AT.

**TABLA IV.** Actitud hacia las clases de ciencias antes y después de la acción en valor medio y desviación típica. \*Se incluye el nivel de significación del contraste de hipótesis.

	Media	Dt	t	gl	p*
ATA	30.15	5.56	5.46	299	<.00
ATD	32.51	5.13			

Por un lado, hay que destacar que de manera general se da una mejora en la actitud hacia las clases de ciencias tras la participación en el proyecto. Dado que la diferencia de medias en la actitud total es estadísticamente significativa, podemos decir que se incrementa la misma en 2.36 tras la participación en el proyecto.

Por otro lado, se realiza un análisis exploratorio de las variables cuantitativas, ATA y ATD, y los posibles factores relacionados con ellas mediante la correlación de Spearman. Se obtiene que los factores asociados con la actitud total son etapa (142,  $p = .01$ ) y zona del colegio (.175;  $p < .00$ ), que correlacionan débilmente con ATA, y un poco más fuertemente con ATD (.43,  $p < .00$ , para etapa; .22;  $p < .00$  para zona de colegio). Con el factor sexo no se obtiene asociación ni para ATA (.05;  $p = .38$ ) ni para ATD (.10;  $p = .06$ ).

Se analiza la normalidad de la distribución de ATA y ATD respecto a los tres factores estudiados: etapa, sexo y zona del colegio. Se emplea el test de Kolmogorov-Smirnov, con el que se obtiene que son normales las distribuciones para la etapa ( $p = .20$ ), y no lo son para sexo ( $p = .04$ ) y zona del colegio ( $p < .01$ ). Por ello, los contrastes de medias que se usan son un ANOVA en el caso de la etapa, y un U de Mann-Whitney para el estudio del sexo y la zona del colegio.

Se detallan a continuación los datos obtenidos para cada uno de los factores analizados.

## Factor Etapa

Los resultados de las medias de la actitud para las etapas estudiadas, primaria, secundaria obligatoria y bachillerato, se muestran en la tabla V.

**TABLA V.** Valor medio y desviación típica de la actitud hacia las clases de ciencias antes y después de la acción para las etapas estudiadas. \*Se incluye el nivel de significación del contraste de hipótesis

ETAPA	ATA		ATD		t	gl	p*
	Media	Dt	Media	Dt			
Primaria	30.62	5.94	33.91	4.41	-6.61	210	<.00
Secundaria	28.71	5.62	28.89	5.47	-.13	66	.90
Bachillerato	29.91	3.55	29.96	4.37	-.04	22	.97

En la actitud hacia las clases de ciencias antes y después de la acción para las etapas estudiadas, se observa que la más positiva está en primaria y la menos positiva en secundaria obligatoria. Por otro lado, al comparar las medias el ANOVA nos indica que existe diferencia con la etapa ( $F = 3.067$ ;  $p=.048$ ) para la actitud antes, ATA, y también para la actitud después ( $F= 44.963$ ;  $p<.00$ ), ATD. Al comparar las diferencias por alumno, el análisis nos muestra que son distintas las diferencias para cada etapa ( $F=5.709$ ;  $p<.00$ ).

Sin embargo, cuando comparamos la actitud antes y después de la acción de fomento de vocaciones científicas para la misma etapa, ATD-ATA, encontramos que existe diferencia de +3.29, significativa estadísticamente en la etapa de primaria únicamente. En las otras dos etapas la diferencia es .13 y .04, casi despreciable, y no tiene significación estadística. Por tanto podemos decir que en la etapa de educación primaria, esta acción tiene un notable efecto positivo sobre la actitud hacia las clases de ciencias de los alumnos, y que no podemos afirmar que exista un efecto similar para las etapas de educación secundaria obligatoria y bachillerato.

## Factor sexo

Los valores de la actitud total para sendos sexos se muestran en la tabla VI.

**TABLA VI.** Actitud media y desviación típica antes y después para los dos sexos. \*Se incluye el nivel de significación del contraste de hipótesis

	Sexo	Media	Dt	U-Mann Whitney*
ATA	Chica	30.45	5.87	p=.38
	Chico	29.84	5.64	
ATD	Chica	32.21	4.87	p=.06
	Chico	32.82	5.40	

La asociación entre ATA y ATD y el sexo no se da en este conjunto de datos, si bien en la actitud después de la intervención podría considerarse marginalmente significativa.

Se profundiza en el análisis del sexo estudiando si pudiera darse alguna diferencia en alguna de las etapas consideradas, dado que se encuentra que los factores sexo y etapa están asociados ( $\chi^2(N=380, 2) = 6.67, p=.04$ ). Se muestran las medias en la tabla VII.

**TABLA VII.** Actitud media y desviación típica antes y después para los dos sexos por etapa. Media (desviación típica)

M(Dt)		Ed. Primaria	ESO	Bachillerato
ATA	Chica	30.97(6.07)	29.45(5.73)	29.11(3.30)
	Chico	30.30(5.82)	27.42(5.27)	30.43(3.74)
ATD	Chica	33.53(4.23)	29.43(5.26)	30.00(3.81)
	Chico	34.28(5.82)	27.96(5.81)	29.93(4.84)

Las medias no son distintas para chicas y para chicos en la etapa de Primaria antes de la intervención ( $U=5432.00, p=.39$ ), ni tampoco en ESO ( $U=427.50, p=.31$ ) o bachillerato ( $U=49.50, p=.40$ ). Algo similar ocurre después de la misma en Primaria ( $U=5317.50, p=.10$ ) y en ESO ( $U=457.50, p=.36$ ). Sin embargo en bachillerato encontramos una diferencia de medias marginalmente significativa ( $U=352.50, p=.07$ ), como ocurre en la muestra total. Esa ligera diferencia implica una actitud ligeramente más positiva en chicos.

## Factor zona del colegio

Los resultados respecto de este factor se muestran en la tabla VIII. Tanto en la actitud antes, ATA, como después, ATD, observamos valores medios diferentes para zonas rurales y urbanas. Estas diferencias tienen significación estadística en ambos casos. Se observa una actitud más positiva en los alumnos de las zonas rurales que en las urbanas en ambos momentos.

TABLA VIII. Actitud media y desviación típica antes y después para las zonas donde se encuentra el colegio. \*Se incluye el nivel de significación del contraste de hipótesis

	Zona colegio	Media	Dt	U-Mann Whitney*
ATA	Rural	31.60	5.54	p<.00
	Urbana	29.41	5.72	
ATD	Rural	33.72	4.57	p<.00
	Urbana	31.15	5.30	

El valor obtenido para la media de ATA en zonas rurales es 31.60, y para ATD 33.72, lo que supone un incremento de 2.12. En las zonas urbanas los valores para ATA y ATD son, respectivamente, 29.41 y 31.15, lo cual significa también un incremento de 1.74, algo menor que en las rurales.

## Discusión

Este trabajo valora el efecto de un programa divulgativo cuya acción principal es desarrollar experimentos adecuados al alumnado participante que implementan los estudiantes, sobre su actitud hacia las clases de ciencias. En la valoración de dicho efecto se tienen en cuenta factores como la etapa educativa, el sexo y el entorno del centro (rural o urbano). Se realiza un estudio transversal con una metodología cuantitativa que emplea un cuestionario breve con una muestra de 380 estudiantes de las etapas de Educación Primaria, Secundaria obligatoria y Bachillerato de los centros participantes en el proyecto. Se emplea

un diseño cuasiexperimental, pues se administra un pretest y postest. Los resultados muestran: una actitud más positiva hacia las clases de ciencias en la escuela en el postest ( $t=5.46$ ;  $gl=299$ ;  $p<.01$ ), siendo esta mejora mayor en el alumnado de la etapa de educación primaria ( $t=-6.61$ ;  $gl=210$ ;  $p<.00$ ); en el alumnado de zonas rurales el efecto del programa provoca un incremento más notable ( $p<.00$ ); y se aprecia una influencia del sexo en el efecto del programa sobre la actitud ( $p=.06$ ) marginalmente significativa en la etapa de bachillerato.

En este estudio, en líneas generales, se observa una actitud más positiva hacia las clases de ciencias tras la participación en el proyecto. Los resultados de mejora de la actitud mediante acciones dentro y fuera de la escuela han sido reportados por otros autores. Los estudios que se han encontrado más comparables con el actual son los trabajos de Weinberg, Basiley Albright (2011), el de Cardona et al (2012), y el de Barmby et al (2008). Los tres proponen acciones externas a la escuela llevadas a cabo por personas diferentes al docente habitual, y estudian su efecto sobre la actitud. Si comparamos el tipo de acción realizada, los más similares al nuestro, son el primero y el tercero. Difieren ambos en que en el primero se valora la motivación hacia las ciencias y las matemáticas, y no la actitud hacia las clases de ciencias, como sí se hace en el tercero. Por lo tanto, centramos nuestra atención en este último por su similitud con este trabajo. Como en el trabajo que se presenta aquí, en el trabajo de Barmby et al (2008) se emplean experimentos itinerantes (Lab in a Lorry) desarrollados por científicos, y se valora su influencia en la actitud hacia las clases de ciencias. En este caso obtienen una mejora de la actitud, en línea con los resultados aquí presentados. Sin embargo, difieren ambos estudios en que en aquel se trabajó solo con alumnos de enseñanza secundaria, mientras que en este caso se reportan datos de las etapas de educación primaria, secundaria obligatoria y bachillerato. Además, aquí se obtiene una mejora de la actitud mayor justo en el alumnado de la etapa primaria. Por lo tanto, los datos referentes a la etapa de primaria y a la mejora obtenida con esta acción constituyen la parte más destacable de este trabajo, dado la ausencia de estudios previos que contemplen acciones que involucren experimentación de los estudiantes con alumnado de esta etapa educativa. Podrían justificarse las diferencias observadas en las diferentes etapas en que el tipo de intervención no afecta por igual a los distintos grupos, bien sea por el grado de participación de los alumnos, por el ajuste a sus conocimientos

o expectativas previas, o por alguna variable que no se ha considerado en este estudio.

No obstante, a pesar de que en este trabajo se afirma no solo que la actitud es más positiva tras la intervención, sino que esta tiene una influencia más positiva en los estudiantes de la etapa de educación primaria, somos conscientes de que tiene limitaciones metodológicas: la muestra está irregularmente distribuida por etapas, y no representativa de la población estudiantil del territorio nacional. Esto ocurre porque la composición muestral está condicionada por la desigual participación en el proyecto de los centros, lo que limita las implicaciones del estudio.

Respecto del sexo, no se observa a nivel general influencia del sexo en la actitud hacia las clases de ciencias ni antes ni después de la intervención. Se destaca una diferencia de medias marginalmente significativa en los estudiantes de bachillerato. Los resultados encontrados en la bibliografía a este respecto son controvertidos. Por un lado, en estudios como el de Weinberg, Basiley Albright (2011), y el realizado por (Fernández Cézar & Solano Pinto, 2017) no se reportan diferencias, en contraposición con lo que se muestra en el trabajo de Marbà-Tallada y Solsona-Pairó (2012) o el de Smith, Pasero y McKenna (2014), que informan de que las chicas presentan una actitud hacia las ciencias más negativa, y destacan la diferencia de preferencias científicas entre chicos y chicas, siendo los primeros los que muestran más inclinación hacia tareas científicas. Por otro lado, a pesar de que en el trabajo de Fernández Cézar y Solano Pinto (2017) no se encuentra diferencia por sexos, se reporta que entre aquellos estudiantes de educación primaria que muestran una más positiva actitud hacia las clases de ciencias, la mayoría son chicas. La diferencia marginalmente significativa encontrada en este estudio en los alumnos de bachillerato está en línea con los trabajos de Marbà-Tallada y Solsona-Pairó (2012) y el de Smith, Pasero y McKenna (2014), pues se encuentra una actitud más negativa y un menor efecto en las chicas, mientras que está en línea con los estudios de Fernández Cézar y Solano Pinto (2017), y Weinberg, Basiley Albright (2011) en estudiantes de primaria y secundaria. Podríamos argumentarlo en base a la diferente edad de los estudiantes, de manera que la ausencia de diferencia en los alumnos de menor edad puede achacarse a que tengan menos asumidos los roles que socialmente vinculan las ciencias al sexo masculino. Por otro lado, apoyaría la actitud positiva en las chicas el hecho de que, aunque es conocido el declive general de vocaciones científicas desde

la primera década de este siglo (Vázquez Alonso, & Manassero–Mas, 2008), actualmente los números en las facultades de bioquímica o medioambiente indican una mayoría de alumnado femenino. No ocurre esto en carreras técnicas como en algunas ingenierías. Por lo tanto, además de la diferencia de edad de los estudiantes, identificamos que otra limitación de este trabajo proviene de que el instrumento empleado para cuantificar la actitud hacia las clases de ciencias en este trabajo no permite diferenciar entre ciencias de la salud o técnicas. Este sería un aspecto a explorar en el futuro para ver si hay sesgo de género en todas o algunas disciplinas científicas.

Respecto de la zona donde se ubica el colegio, rural o urbana, se encuentran pocos estudios que analicen este aspecto de manera general. El de Chetcuti y Kioko (2012) muestra que las chicas de entornos rurales tienen actitudes hacia las ciencias más positivas que las de entornos urbanos en Kenya. Este aspecto no ha sido estudiado en contextos más cercanos al nuestro. En este estudio los resultados coinciden con los de Chetcuti y Kioko en tanto que el alumnado de zonas rurales tiene actitud más positiva que los de zonas urbanas, tanto antes como después de la participación en el proyecto. En el trabajo de Fernández César y Solano Pinto (2017) se reporta también que los alumnos de zonas rurales tienen una actitud hacia las clases de ciencias más positiva. Dado la diversidad de oferta extraescolar entre los entornos rurales y urbanos, son necesarias acciones divulgativas como este proyecto para fomentar la igualdad de oportunidades. También son necesarios más estudios en esta línea que saquen a la luz la situación educativa del alumnado en entornos rurales en España y Europa.

No obstante, dada la no representatividad de la muestra, que no está ponderada en el aspecto rural/urbano como lo está la población en el territorio nacional, como se ha señalado al inicio de esta discusión, es necesaria una investigación más profunda para confirmar (o no) si en las zonas rurales el alumnado presenta una actitud hacia las ciencias más positiva de forma generalizada.

## **Conclusiones: limitaciones y prospectiva**

Con todo esto, concluimos que la valoración de intervenciones que involucran la experimentación implementada con los estudiantes,

acompañados por investigadores de primera línea, tiene una influencia positiva en los estudiantes de la etapa de primaria, y en menor medida en los estudiantes de ESO o Bachillerato. El mayor descenso se produce entre la educación primaria y la secundaria, tanto antes como después de la intervención. Hay algunos trabajos recientes que analizan las emociones tanto del profesorado como del alumnado de esta etapa educativa (Cortés, 2015), y recalcan la influencia que tiene en los siguientes itinerarios seguidos por los estudiantes. No encontramos asociación de la AT con el sexo con carácter general, y si se asocia con la zona donde se ubica el colegio, siendo esta más positiva en los estudiantes de las zonas rurales. Se produce un aumento tras la participación en el proyecto en ambas zonas geográficas.

Dadas las implicaciones que este estudio plantea, siendo conscientes de sus limitaciones debidas tanto a la composición de la muestra como al tipo y duración de la intervención desarrollada, sería necesario ahondar en el estudio de la actitud hacia las clases de ciencias y los tipos de intervenciones que pudieran mejorarlas, sobre todo en estudiantes de ESO y bachillerato. Esta intervención tiene un aspecto positivo y es que involucra a los estudiantes en la experimentación, lo cual contribuye de manera positiva en la actitud hacia las clases de ciencias. El aspecto negativo, o punto débil, es su corta duración, que hace que sus efectos sean leves y se diluyan con el tiempo. Por eso es menos aconsejable para actuar sobre alumnado de la educación secundaria en general, que por su edad posee ideas más afianzadas y ya menos modificables, y requiere de intervenciones mantenidas en el tiempo. Por lo tanto, creemos que para conseguir valores altos de la actitud hacia las clases de ciencias desde la educación primaria y que se mantengan en las etapas posteriores, potenciar el desarrollo cognitivo y emocional del alumno e incidir sobre los futuros itinerarios científicos a seguir por los estudiantes, son necesarias acciones que involucren a los alumnos de manera activa mediante la experimentación, que les permitan manipular y no solo observar las experiencias, tomar decisiones en vez de simplemente ver sus efectos, similares a la presentada aquí, pero no puntual y limitada en el tiempo sino con carácter más duradero o incluso permanente. Para ello, lo más efectivo sería dirigir las acciones hacia los maestros y profesores de materias científicas. En el desarrollo de este proyecto venimos observando en las visitas a los centros que muchos maestros sienten inseguridad para incorporar los experimentos como base de sus

clases de ciencias. Por ello, sería necesaria la formación adecuada y la elaboración de materiales y recursos didácticos que pudieran ayudarles en esta tarea, con el fin de que la experimentación con la participación de los alumnos constituyera la parte fundamental de sus clases de ciencias.

## Referencias bibliográficas

- Abad, F. J., Olea, J., Ponsoda, V. & García, C. (2011). *Medición en ciencias sociales y de la salud*. Madrid: Síntesis.
- Aikenhead, G. S., & Ryan, A. G. (1992). The Development of a New Instrument: 'Views on Science—Technology—Society' (VOSTS). *Science Education*, 76(5), 477-491.
- Alonso, Á. V., & Mas, M. A. M. (2007). *Los intereses curriculares en ciencia y tecnología de los estudiantes de secundaria*. Universitat Illes Balears, Servei de Publicacions i Intercanvi Científic.
- Alonso, Á. V., & Mas, M. A. M. (2008). La elección de asignaturas de ciencias: análisis de los factores determinantes. *Revista española de Pedagogía*, 541-558.
- Alonso, Á. V., & Mas, M. A. M. (2009). La relevancia de la educación científica: actitudes y valores de los estudiantes relacionados con la ciencia y la tecnología. *Enseñanza de las ciencias: revista de investigación y experiencias didácticas*, 27(1), 33-48.
- Alsop, S. (Ed.). (2005). *Beyond Cartesian Dualism: Encountering Affect in the Teaching and Learning of Science* (Vol. 29). Springer Science & Business Media.
- Barmby, P., Kind, P. M., & Jones, K. (2008). Examining changing attitudes in secondary school science. *International journal of science education*, 30(8), 1075-1093.
- Bennássar, A., Vázquez, A., Manassero, M. A. y García-Carmona, A. (Coords.) (2010). *Ciencia, tecnología y sociedad en Iberoamérica: una evaluación de la comprensión de la naturaleza de ciencia y tecnología*. Madrid, España: Organización de Estados Iberoamericanos para la Educación, la Ciencia y la Cultura (OEI).
- Bethke Wendell, K., & Rogers, C. (2013). Engineering design-based science, science content performance, and science attitudes in elementary school. *Journal of Engineering Education*, 102(4), 513-540.

- Briñol, P., Horcajo, J., De la Corte, L., Valle, C., Gallardo, I. y Díaz, D. (2004). El efecto de la ambivalencia evaluativa sobre el cambio de actitudes. *Psicothema*, 16, 373-377.
- Cardona, L. Z., Angulo, F., Soto, C. A., Quintero, S. M., Ceballos, A. F., Delgado, E., & Cifuentes, L. J. (2012). ¿Contribuyen los talleres en el Museo de Ciencias a fomentar actitudes hacia la conservación del ambiente? *Enseñanza de las ciencias: revista de investigación y experiencias didácticas*, 30(3), 53-70.
- Chen, S. (2006). Development of an instrument to assess views on nature of science and attitudes toward teaching science (VOSE). *Science Education*, 90(5), 803-819.
- Chetcuti, D. A., & Kioko, B. (2012). Girls' attitudes towards science in Kenya. *International Journal of Science Education*, 34(10), 1571-1589.
- Cortés, A. B. B. (2015). Las emociones en la enseñanza y el aprendizaje de las ciencias en educación secundaria. *Enseñanza de las ciencias: revista de investigación y experiencias didácticas*, 33(3), 199-200.
- de Pro Bueno, A., & Pérez Manzano, A. (2014). Actitudes de los alumnos de Primaria y Secundaria ante la visión dicotómica de la Ciencia. *Enseñanza de las Ciencias*, 32(3), 111-132.
- Di Martino, P., & Sabena, C. (2010). Teachers' beliefs: the problem of inconsistency with practice. In *Proceedings of the XXXIV Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 313-320).
- Di Martino, P. & Zan, R. (2010). Me and maths: towards a definition of attitude grounded on students' narratives. *Journal of Mathematics Teacher Education*, 13(1), 27-48.
- Eiser, J.R. (1989). *Psicología social. Actitudes, cognición y conducta*. Madrid: Pirámide.
- Fernández Cézar, R., & Solano Pinto, N. (2017). Actitud hacia las clases de Ciencias Naturales en la educación primaria en España. *Revista electrónica de investigación educativa*, 19(4), 112-123. doi: <https://doi.org/10.24320/redie.2017.19.4.1393>
- Heckman, J., Stixrud, J. y Urzua, S. (2006). The effects of cognitive and non- cognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics*, 24 (3), 411-482.
- Hong, Z. R., Lin, H. S., Wang, H. H., Chen, H. T., & Yang, K. K. (2013). Promoting and scaffolding elementary school students' attitudes

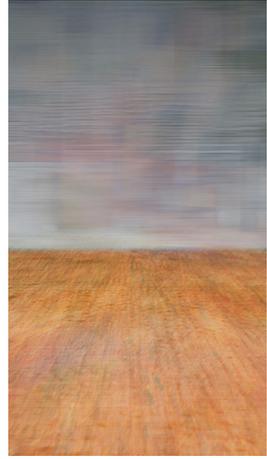
- toward science and argumentation through a science and society intervention. *International Journal of Science Education*, 35(10), 1625-1648.
- Kennedy, J., Quinn, F., & Taylor, N. (2016). The school science attitude survey: a new instrument for measuring attitudes towards school science. *International Journal of Research & Method in Education*, 39(4), 422-445. doi: 10.1080/1743727X.2016.1160046
- Manassero, M. A., Vázquez, A., & Acevedo, J. A. (2003). Cuestionario de opiniones sobre ciencia, tecnología y sociedad (COCTS). *Princeton, NJ: Educational Testing Service. Información en línea en <http://www.ets.org/testcoll>*.
- Manassero-Mas, M. A., Vázquez-Alonso, Á., & Acevedo-Díaz, J. A. (2001). *Cuestionario de Opiniones Sobre Ciencia Tecnología Y Sociedad (COCTS)*. Universidad de las Islas Baleares, Departamento de Psicología.
- Marbà-Tallada, A. & Márquez Bargalló, C., (2010) ¿Qué opinan los estudiantes de las clases de ciencias? Un estudio transversal de sexto de Primaria a cuarto de ESO. *Enseñanza de las Ciencias*, 28(1), 19–30.
- Marbà-Tallada, A., & Solsona-Pairó, N. (2012). Identificación e interpretación de las posibles desigualdades formativas en ciencias de chicos y chicas en la educación obligatoria y el bachillerato. *Cultura y Educación*, 24(3), 289-303. Si hay diferencia
- Mellado, V., Blanco, L. J., Borrachero, A. B., & Cárdenas, J. A. (2013). *Las emociones en la enseñanza y el aprendizaje de las ciencias y las matemáticas*. Ed: DEPROFE. Badajoz, España.
- Potvin, P. & Hasni, A. (2014). Interest, motivation and attitude towards science and technology at K-12 levels: a systematic review of 12 years of educational research. *Studies in Science Education*, 50 (1) 85–129, <http://dx.doi.org/10.1080/03057267.2014.881626>
- Osborne, J., Simon, S. y Collins, S. (2003). Attitudes towards science: a review of the literature and its implications. *International Journal of Science Education*, 25(9), 1049 – 1079.
- ROSE, A. C. D. P. (2005). Proyecto ROSE: relevancia de la educación científica. *Revista Eureka sobre enseñanza y divulgación de las ciencias*, 2(3), 440-447
- Rubba, P. A., & Harkness, W. L. (1993). Examination of preservice and in-service secondary science teachers' beliefs about science-technology-society interactions. *Science Education*, 77(4), 407-431.

- Saleh, I. M., & Khine, M. S. (Eds.). (2011). *Attitude research in science education: Classic and contemporary measurements*. IAP.
- Schreiner, C. & Sjøberg, S. (2004). Showing the seeds of ROSE. Background, Rationale, Questionnaire Development and Data Collection for ROSE (The Relevance of Science Education) - a comparative study of students' views of science and science education. *Acta Didactica*, (4/2004). Published: Dept. of Teacher Education and School Development, University of Oslo, Norway. Retrieved on 2/12/2014 from <http://www.ils.uio.no/forskning/rose/documents/AD0404.pdf>.
- Siegel, M. A., & Ranney, M. A. (2003). Developing the changes in attitude about the relevance of science (CARS) questionnaire and assessing two high school science classes. *Journal of Research in Science Teaching*, 40(8), 757-775.
- Smith, T. J., Pasero, S. L., & McKenna, C. M. (2014). Gender effects on student attitude toward science. *Bulletin of Science, Technology & Society*, 34(1-2), 7-12.
- Unal, A., & Altıparmak, M. (2013). The Effects of Webquests on Elementary School Seventh Grade Students' Science Achievement, Attitude towards Science and Attitude towards Web Based Education while Teaching Science Subjects. In *EDULEARN13 Proceedings* (pp. 4165-4173). IATED.
- Vázquez, A. (2013). La educación científica y los factores afectivos relacionados con la ciencia y la tecnología. En *Las emociones en la enseñanza y el aprendizaje de las ciencias y las matemáticas*, (pp.245-278). Ed: DEPROFE. Badajoz, España.
- Vázquez Alonso, Á., Acevedo Díaz, J. A., Manassero Mas, M. A., & Acevedo Romero, P. (2006). Actitudes del alumnado sobre ciencia, tecnología y sociedad, evaluadas con un modelo de respuesta múltiple. *Revista Electrónica de Investigación Educativa*, 8(2), 1-37.
- Vázquez, Á., & Manassero, M. A. (2008). El declive de las actitudes hacia la ciencia de los estudiantes: un indicador inquietante para la educación científica. *Revista Eureka sobre Enseñanza y Divulgación de las ciencias*, 5(3).
- Vázquez- Alonso, A. & Manassero- Mas, M.A. (2011). El descenso de las actitudes en la educación obligatoria. *Ciência & Educação*, 17(2), 249-268.
- Vázquez- Bernal, B. y Jiménez- Pérez, R. (2016). Modelización de un constructo teórico sobre la percepción del alumnado en procesos de indagación en ciencias. *Revista de Psicodidáctica*, 21(1), 25-44.

Weinberg, A. E., Basile, C. G., & Albright, L. (2011). The effect of an experiential learning program on middle school students' motivation toward mathematics and science. *RMLE Online*, 35(3), 1-12.

**Información de contacto:** Raquel Fernández-César. Universidad de Castilla la Mancha, Fac. de Educación de Toledo, Dpto. de Matemáticas. Área de Didáctica. Av. Carlos III s/n, 45071, Toledo, Spain. E-mail: raquel.fcezar@uclm.es





## **Reseñas**



**Pérez Pérez, C. (2016). *Educación en valores para la ciudadanía. Estrategias y técnicas de aprendizaje*. Bilbao: Desclée de Brouwer.**

En la medida que el ser humano está liberado del instinto, tiene capacidad moral y la posibilidad de hacerse a sí mismo a través de la educación. Por esta razón los valores cívicos y ciudadanos constituyen un contenido fundamental sin el cual la educación no sería hoy posible.

En la obra reseñada, el profesor Cruz Pérez presenta la ciudadanía como un concepto polifacético que abarca diversos ámbitos como el legal, el social, el económico, el cívico, el intercultural o el cosmopolita, cada uno de ellos con sus implicaciones educativas, amplificadas durante las últimas décadas por el creciente interés que ha suscitado el tema en el contexto de la Unión Europea.

Si nos ceñimos al ámbito del Sistema Educativo Formal y tal como señala el autor, el profesorado encuentra serias dificultades para abordar la educación de los valores, entre las que se destaca: la falta de tradición para programar este tipo de contenidos en la enseñanza, la dificultad de emplear estrategias para la modificación de valores y actitudes, la complejidad de introducir la red de organización social y familiar en la escuela, el peligro del adoctrinamiento, la influencia de los Medios de Comunicación Social, la férrea disposición de muchos profesores a enseñar en solitario; y sus carencias formativas para ejercer una enseñanza de calidad en valores.

En cualquier caso, la pregunta clave que se plantea es “¿en qué valores educar?”. Respondiendo a esta cuestión, en el libro se hace una apuesta clara por la educación en una ética de mínimos, en sintonía con las propuestas de autores relevantes como A. Cortina o J. Escámez. Bajo este concepto se agrupan diversas propuestas basadas en la idea de que en una sociedad abierta y plural es necesario e inevitable educar en unos valores que podamos asumir entre todos.

Para abordar estas cuestiones clave, en el libro se describen las principales teorías desde las que es posible educar en valores: sociologismo, clarificación de valores, desarrollo del juicio moral, educación del carácter y filosofía para niños, a la vez que se presentan los planteamientos pedagógicos que se derivan de cada una de ellas.

La educación en valores para el ejercicio de la ciudadanía ocupa un capítulo completo, donde se estudia el recorrido histórico, la globalización y emersión de escenarios colmados de pluralidad cultural, los aspectos legales como derechos normativos, el contexto europeo y la actual sociedad de la información y la comunicación. También se presentan las estrategias que el autor considera más adecuadas para educar en la ciudadanía en el contexto escolar, entre las que cabe destacar la organización de los centros según un modelo de participación democrática, la educación para la convivencia y resolución de conflictos, así como la cultura de colaboración y solidaridad con las necesidades de la comunidad, entre otras.

También se dedica un capítulo a la descripción y análisis de las técnicas para el aprendizaje de valores que han resultado más eficaces para la formación de buenos ciudadanos. Para ello se seleccionan dieciséis técnicas concretas en las que se presenta una detallada descripción de los objetivos, contenidos, aplicación, función del profesor y metodología de evaluación. Con las mismas se pretende proporcionar herramientas de intervención educativa al profesorado, en un ámbito en el que existe una gran carencia de recursos e información.

Las técnicas finalizan generalmente con el debate e intercambio de impresiones que han surgido en el transcurso de la actividad, otorgando un cariz conclusivo al trabajo realizado. En última instancia, se trata de llevar valores a la acción y a la reflexión de una manera constantemente propositiva y compartida.

El libro pone su broche final abordando el complejo tema de la evaluación de los valores, entendida ésta como continua, global, formativa y procesual, y en la que cobran especial relevancia elementos como la distinción entre la opinión personal y los conocimientos y habilidades del alumnado, la percepción de comentarios, la adquisición de valores morales y el uso de instrumentos y técnicas de evaluación de diferente tipo.

Se trata, en suma, de un libro que recoge las actualizaciones teóricas sin renunciar a los planteamientos más clásicos ni tampoco a prestar

soluciones prácticas para la educación en valores, tanto desde un prisma escolar como desde una creciente preocupación social y ciudadana.

*Fran J. Garcia-Garcia*

**Monarca, H. y Thoilliez, B. (Coords.) (2017). *La profesionalización docente: debates y propuestas*. Madrid: Síntesis. pp. 146. ISBN: 978-84-9171-045-5**

A lo largo de sus diez capítulos, los autores nos hacen llegar la preocupación existente por «llenar nuestras aulas de “docentes eficaces”» (p.9), y a la vez que tratan de sensibilizarnos ante el debilitamiento actual de la profesión docente respaldándose en una literatura revisada, defienden la importancia de la formación inicial del profesorado.

Enric Prats y Ana Marín apuestan por la universidad como principal formadora de futuros docentes, pero invitan a reflexionar sobre su papel insistiendo en que ha de ofrecer una educación sólida basada en una triple dimensión formadora: cultural, intelectual e investigadora.

Paul Standish critica la idea de educación como intercambio dentro de una *economía cerrada*, donde la misma genera unos resultados y se pone al servicio de una imperante rendición de cuentas en la práctica educativa. Por ello, defiende unas «economías del exceso de la alteridad y de la intensidad» (p.40) para entender qué es lo verdaderamente importante.

Después, Héctor Monarca, desarrolla la idea de profesionalismo docente y saca a la luz algunas fragmentaciones y contradicciones. Destacando su carácter polisémico, defiende que pueden darse formas alternativas que deriven en «praxis transformadoras y liberadoras» (p.50).

Bianca Thoilliez defiende dos ideas principales: por un lado, que orientar la profesión docente a producir resultados objetivos de aprendizaje agrieta los fundamentos del acto de educar; y por otro, que la mejor manera de preparar a los futuros docentes no reside en un modelo de desarrollo profesional docente (DPD) basado en evidencias, pues educar es «más una cirugía heroica que un protocolo técnico» (p.62).

Más tarde, Geo Saura y Noelia Fernández-González hablan de “docentes gobernados por los números” y de reformas educativas contemporáneas basadas en lógicas mercantiles, abordando el DPD como un nuevo profesionalismo pretendido donde los gobiernos están modelando la profesión. Finalmente, critican la aparición de mecanismos de endoprivatización que están desprofesionalizando la docencia a nivel global.

Por otro lado, Fernando Gil Cantero nos habla de los maestros como personas que deben cultivarse a sí mismas y preocuparse por su propio crecimiento personal e intelectual, en otras palabras, que quieren ser mejores. Para ello, reclama un dominio de la teoría pedagógica, *más reflexión y juicio educativo para dotar de consistencia los propios ideales y convicciones pedagógicas*.

Conceptualizar la identidad docente es tarea de Tania Alonso, para quien la formación inicial es clave para asentar las bases de la identidad profesional docente y superar la *crisis de identidad* o «la falta de voz propia del docente a la hora de decirse a él mismo quién es y qué papel tiene en la escuela y en la sociedad» (p.93). Apoyada en la literatura de Charles Taylor, reivindica la necesidad de que exista una identidad docente sólida.

Jesús Manso compara la importancia de la formación inicial del profesorado para el DPD con los primeros años de vida para el desarrollo de las personas. Además, define el periodo de inducción a la profesión docente y reflexiona sobre las diversas rupturas que surgen en este proceso. Concluye pidiendo un mayor protagonismo del docente en las medidas que se lleven a cabo en torno a estos debates.

David Reyero nos recuerda que «es imposible ser un buen profesor si no se domina aquello que se enseña» (p.125), y demanda una formación disciplinar sólida. Critica el plano secundario en el que parecen encontrarse los contenidos culturales y propone revalorizarlos en el currículum escolar, revisar las competencias pedagógicas, reclutar a los mejores docentes y formarlos en esos contenidos culturales.

Finalmente, Inmaculada Egido analiza el prácticum docente tras revisar narrativamente diversas investigaciones. Centrándose en el ámbito nacional, recoge los puntos clave en relación con las funciones y tareas realizadas por los estudiantes en prácticas, los tutores y supervisores del prácticum, los centros de prácticas, la colaboración entre estos y la universidad, y la organización, desarrollo y evaluación de este periodo.

Las conclusiones sugeridas, nos dice, podrían ser un primer paso para el futuro diseño de un prácticum de calidad.

El libro recoge una revisión completa de los principales puntos que definen el DPD, y a lo largo de sus *páginas* los autores consiguen atraer nuestro interés invitándonos a formar parte de los debates educativos que se presentan y generando una inevitable autorreflexión. Algunos autores exigen a los futuros docentes más reflexión, contacto y discusión de la literatura, ¿por qué no empezar por dejar caer en sus mesas este libro? *La profesionalización docente: debates y propuestas* debería pasar por sus manos y ser fuente de debate en las aulas universitarias de las facultades de educación.

***Virginia Cabadas Maldonado***



La *Revista de Educación* es una publicación científica del Ministerio de Educación, Cultura y Deporte español. Fundada en 1940, y manteniendo 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 en este campo, tanto desde una perspectiva nacional como internacional. La revista es editada por la Subdirección General de Documentación y Publicaciones, y actualmente está adscrita al Instituto Nacional de Evaluación Educativa de la Dirección General de Evaluación y Cooperación Territorial.



NIPO línea: 030-15-016-X  
NIPO ibd: 030-15-017-5  
ISSN línea: 1988-592X 0034-8082  
ISSN papel: 0034-8082

[www.mecd.gob.es/revista-de-educacion](http://www.mecd.gob.es/revista-de-educacion)