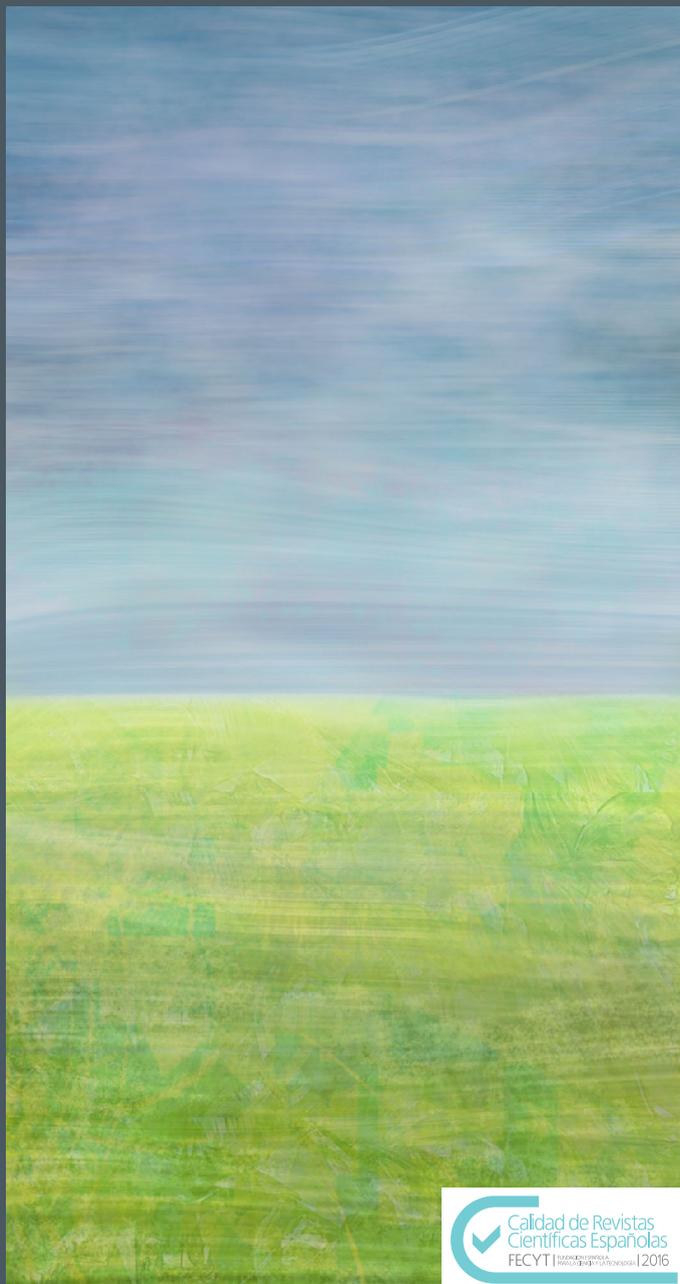


revista de **e**EDUCACIÓN

Nº 383 ENERO-MARZO 2019



GOBIERNO
DE ESPAÑA

MINISTERIO
DE EDUCACIÓN
Y FORMACIÓN PROFESIONAL



revista de
eEDUCACIÓN



Nº 383 ENERO-MARZO 2019

revista de EDUCACIÓN

Nº 383 Enero-Marzo 2019

Revista trimestral

Fecha de inicio: 1952



MINISTERIO DE EDUCACIÓN Y FORMACIÓN PROFESIONAL
SECRETARÍA DE ESTADO DE EDUCACIÓN Y FORMACIÓN PROFESIONAL

Instituto Nacional de Evaluación Educativa
Paseo del Prado, 28, 4.ª planta
28014 Madrid
España

Edita
© SECRETARÍA GENERAL TÉCNICA
Subdirección General de Atención al Ciudadano, Documentación y Publicaciones

Catálogo de publicaciones del Ministerio: mecd.gob.es
Catálogo general de publicaciones oficiales: 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

Alejandro Tiana Ferrer
Secretario de Estado de Educación y Formación Profesional

VOCALES

Fernando Gurrea Casamayor
Subsecretario de Educación y Formación Profesional

Consuelo Vélaz de Medrano Ureta
Directora General de Evaluación y Cooperación Territorial

Clara Sanz López
Directora General de Formación Profesional

Diego Fernández Alberdi
Director General de Planificación y Gestión Educativa

Liborio López García
Secretario General Técnico

Carmen Tovar Sánchez
Directora del Instituto Nacional de Evaluación Educativa

Amparo Barbolla Granda
Subdirectora General de Atención al Ciudadano,
Documentación y Publicaciones

Nuria Manzano Soto
Directora del Centro Nacional de Innovación e Investigación Educativa

Montserrat Grañeras Pastrana
Subdirectora General de Ordenación Académica.
Responsable de la Unidad de Igualdad

CONSEJO EDITORIAL

DIRECTORA

Carmen Tovar Sánchez

EDITOR JEFE

José Luis Gaviria Soto

EDITOR ADJUNTO

David Reyero García

VOCALES

Antonio Cabrales Goitia (University College London); Caterina Casalmiglia (Universitat Autònoma de Barcelona); Antonio Lafuente García (Consejo Superior de Investigaciones Científicas); Leoncio López-Ocón Cabrera (Consejo Superior de Investigaciones Científicas); Carlos Marcelo García (Universidad de Sevilla); Miquel Martínez Martín (Universitat de Barcelona); Francisco Michavila Pitarch; (Universidad Politécnica de Madrid); Juan Manuel Moreno Olmedilla (Banco Mundial); Clara Eugenia Núñez (Universidad Nacional de Educación a Distancia); Lucrecia Santibáñez (Claremont Graduate University); Denise Vaillant (Universidad ORT, Uruguay); Pablo Zoido (Banco Interamericano de Desarrollo).

REDACCIÓN

Jefe de Redacción: Jorge Mañana Rodríguez

Colaboradores: Ruth Martín Escanilla y Óscar Urrea Ríos

ASESORES CIENTÍFICOS

Internacional

Aaron Benavot (State University of New York, SUNY-Albany); Abdeljalil Akkari (Universidad de Ginebra); Mark Bray (University of Hong Kong); José Joaquín Brunner (Universidad Diego Portales, Chile); Dirk Hastedt (Executive Director, International Association for the Evaluation of Educational Achievement, IEA); Felipe Martínez Rizo (Consejero Técnico del INEE, México); Marie-Hélène Doumet (INES Programme, OCDE); Andreas Schleicher (Director, Directorate for Education and Skills, OCDE).

Nacional

Teresa Aguado Odina (Universidad Nacional de Educación a Distancia); Margarita Bartolomé (Universitat de Barcelona); Antonio Bolívar (Universidad de Granada); Josefina Cambra Giné (Colegio de Doctores y Licenciados); Anna Camps i Mundó (Universitat Autònoma de Barcelona); César Coll Salvador (Universitat de Barcelona); Agustín Dosil Maceira (Universidad LiberQuaré); Gerardo Echeita Sarrionandia (Universidad Autónoma de Madrid); Juan Manuel Escudero Muñoz (Universidad de Murcia); Mariano Fernández Enguita (Universidad Complutense de Madrid); Joaquín Gairín Sallán (Universitat Autònoma de Barcelona); José Luis García Garrido (Universidad Nacional de Educación a Distancia); Daniel Gil Pérez (Universitat de València); Fuensanta Hernández Pina (Universidad de Murcia); Carmen Labrador Herraiz (Universidad Complutense de Madrid); Miguel López Melero (Universidad de Málaga); Elena Martín Ortega (Universidad Autónoma de Madrid); Rosario Martínez Arias (Universidad Complutense de Madrid); Inés Miret (Neturity S.L., Madrid); Juan Ignacio Pozo (Universidad Autónoma de Madrid); Joaquim Prats Cuevas (Universitat de Barcelona); Manuel de Puelles (Universidad Nacional de Educación a Distancia); Tomás Recio Muñoz (Universidad de Cantabria); Luis Rico Romero (Universidad de Granada); Juana M.ª Sancho Gil (Universitat de Barcelona); Mercedes Vico Monteoliva (Universidad de Málaga); Antonio Viñao Frago (Universidad 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/), 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

ROSENDO MARTÍNEZ-RODRÍGUEZ, CARLOS MUÑOZ LABRAÑA y MARÍA SÁNCHEZ-AGUSTÍ: Conocimientos y creencias del profesorado y su vínculo con las finalidades de la enseñanza de la historia reciente. La transición a la democracia en España como caso controvertido.....	11
ANDRÉS SANDOVAL-HERNANDEZ, DAVID RUTKOWSKI, TYLER MATTA y DANIEL MIRANDA: Pensémoslo de nuevo: ¿Podemos comparar las escalas de antecedentes socioeconómicos?	37
LUÍS DOMÍNGUEZ CASTRO y JOSÉ RAMÓN RODRÍGUEZ-LAGO: Educación y diplomacia cultural en la primavera de Europa (1948-1954).....	63
MARÍA-JESÚS FERNÁNDEZ, MANUEL MONTANERO y MANUEL LUCERO: La evaluación de la competencia narrativa en la educación básica	85
JULIA HABA-OSCA, FRANCISCO GONZÁLEZ-SALA y JULIA OSCA-LLUCH: Las revistas de educación a nivel mundial: un análisis de las publicaciones incluidas en el Journal Citation Reports (JCR) del 2016.....	113
JOSÉ JAVIER VERDUGO-PERONA, JOAN JOSEP SOLAZ-PORTOLÉS y VICENTE SANJOSÉ: Evaluación del Conocimiento Científico en Maestros en formación inicial: el caso de la Comunidad Valenciana.....	133
Reseñas	163



Investigaciones

Conocimientos y creencias del profesorado y su vínculo con las finalidades de la enseñanza de la historia reciente. La transición a la democracia en España como caso controvertido¹

Teacher knowledge and beliefs and their relationship with the purpose of teaching recent history. The transition to democracy in Spain as a controversial case

DOI: 10.4438/1988-592X-RE-2019-383-399

Rosendo Martínez-Rodríguez

Universidad de Valladolid

Carlos Muñoz Labraña

Universidad de Concepción

María Sánchez-Agustí

Universidad de Valladolid

Resumen

En el presente estudio se aborda la vinculación existente entre los conocimientos y creencias del profesorado de historia y sus finalidades y propósitos de enseñanza, sobre un periodo controvertido de la historia reciente: la transición de la dictadura a la democracia en España. Para ello, se ha desarrollado una investigación cualitativa a través de 39 entrevistas a profesores de enseñanza secundaria, agrupando sus interpretaciones en tres clases de pensamiento: *positivo-mítico*, *positivo-crítico* y *negativo-crítico*. Los resultados muestran que la mayoría

⁽¹⁾ Esta investigación forma parte de un proyecto de investigación financiado por el Ministerio de Economía y Competitividad en su Plan Nacional de Investigación (2010-2014), Reference: EDU 2009-09775.

de los profesores expresaron algún tipo de interpretación crítica desarrollando un pensamiento complejo de argumentación y análisis histórico. La variable edad no arrojó resultados destacables, pero sí se observaron diferencias respecto a la variable territorial, predominando la visión *negativa-crítica* en regiones donde se desarrollan movimientos independentistas. No obstante, en relación con las finalidades de la enseñanza de la historia, encontramos una importante distancia entre el discurso ideal y los ejemplos concretos con los que argumentan su posición. Es decir muchos de los profesores con una concepción de la historia crítica y conectada con el presente, encontraban el valor de su enseñanza en un relato ejemplar y cerrado del pasado, pocas veces interpretativo o crítico.

Palabras clave: enseñanza de la historia; creencias del profesor; conocimientos del profesor; finalidades y valores de enseñanza; Transición.

Abstract

This paper considers the link between the knowledge and beliefs of history teachers and their views on the aims and purpose of education, with reference to a controversial period in recent history: the transition from dictatorship to democracy in Spain. For this study, qualitative research was undertaken using interviews with 39 secondary school teachers, grouping their interpretations into three types of thinking: positive-mythical, positive-critical and negative-critical. The results show that most teachers expressed some kind of critical interpretation by developing a complex thinking of historical argumentation and analysis. The age variable age did not show remarkable results, but differences were observed with respect to the geographical variable, predominating the critical-negative view in regions where independence movements are developed. However, in relation to the purposes of history teaching, we find an important distance between the ideal discourse and the concrete examples with which they argue their position. That is, many of the teachers with a conception of critical history and connected with the present, found the value of their teaching in an exemplary and closed story of the past, rarely interpretive or critical.

Keywords: history education; teacher beliefs; teacher knowledge; educational aims and values

Introducción

Las investigaciones sobre el pensamiento del profesorado tienen ya una larga tradición dentro de la investigación educativa. El interés por

conocer el pensamiento docente parte de la idea de que su conducta profesional puede verse condicionada por sus procesos de pensamiento, particularmente sus conocimientos y creencias (Fives y Buehl, 2016). Han sido muchos los estudios que han indagado en los conocimientos y creencias del profesor, desde variados enfoques metodológicos (; Delaney, 2012; Fives y Buehl, 2016; Grossman, 1990; Kagan, 1992; Shulman, 1987).

Además del conocimiento concreto del docente, otros estudios han puesto la atención en la parte más subjetiva del pensamiento, considerando que existe toda una trama ideológica de teorías y creencias que terminan interfiriendo en el pensamiento y finalmente en la práctica pedagógica (Gimeno y Pérez, 1988). Podemos definir las creencias como una forma personal de conocimiento, aquellos supuestos implícitos del profesor sobre los estudiantes, la enseñanza, las clases o la materia (Kagan, 1992).

Nuestra investigación versa sobre los conocimientos y creencias del profesorado de historia sobre su materia, y sobre cómo éstos interactúan con sus fines y propósitos de enseñanza. Todo esto en relación a una parte concreta del conocimiento histórico: la transición de la dictadura franquista a la democracia. Un periodo que forma parte del actual debate historiográfico, social y político en España y que tiene una serie de características que le confieren una relación controvertida con la actualidad.

La historia es una ciencia social en continuo debate, no sólo académico sino también institucional y social, con un importante uso público y político de sus descubrimientos y avances que, así mismo, termina interfiriendo en su enseñanza (Carretero, Berger y Grever, 2017). Estudiar el conocimiento docente sobre un periodo particular como la transición nos permite profundizar en este entramado. Porque desde el conocimiento historiográfico profundo de este periodo, y de sus debates actuales, podremos establecer una clasificación coherente de las interpretaciones y creencias de los profesores, y de ahí determinar la relación que pueda existir con sus finalidades y propósitos educativos.

Por lo tanto, el presente trabajo ha tenido como objetivo analizar los conocimientos y creencias que el profesorado de historia tiene sobre un periodo controvertido de la historia reciente y su vinculación con las finalidades y propósitos de su enseñanza. Nuestras preguntas de investigación han sido:

- ¿Qué conocimientos y creencias tienen los profesores de historia sobre la transición de la dictadura a la democracia en España?

- ¿Qué tipo de interpretación hacen de este periodo histórico controvertido?
- ¿Qué finalidades y propósitos tienen los profesores de historia a la hora de enseñar este periodo histórico?
- ¿Cómo influyen los conocimientos, creencias e interpretaciones del profesorado sobre este caso de historia en su concepción de las finalidades y propósitos de enseñanza?

Fundamentación teórica

Concepciones y creencias sobre la historia

Aunque no existen estudios sobre el conocimiento docente de la transición a la democracia, sí se han realizado numerosas investigaciones sobre las concepciones que los profesores tienen de la historia, como disciplina, y sobre cómo estas interactúan con su enseñanza.

Pioneros en este campo fueron los trabajos de Evans, que estudió las concepciones que los docentes tienen sobre la historia y sus significados (Evans, 1990), estableciendo cinco categorías de concepciones de la historia que caracterizaban cinco tipos de profesor de historia: *storyteller*, *scientific historian*, *relativist/reformer*, *cosmic philosopher*, and *eclectic*. Otros estudios han determinado la relación existente entre las concepciones que se tienen de la historia, desde un punto de vista epistemológico, y el tipo de razonamiento histórico de los profesores (Virta, 2001; Wineburg, 1991). En casos concretos, por ejemplo, se comprobó que los docentes analizaban los documentos históricos de una forma más elemental que la propia disciplina histórica (Bohan y David, 1998; Yeager y Davis, 1996), lo que podría estar en relación con la dificultad de sus estudiantes para alcanzar un adecuado pensamiento histórico.

En una línea más cercana a nuestra investigación, están los estudios que han clasificado el pensamiento del profesor según el tipo de razonamiento histórico que realizan. En este sentido, Yilmaz (2010), atendiendo a una larga tradición de debates historiográficos sobre la naturaleza de la historia, proponiendo categorías de concepciones de la historia: “history as the past” and “history as an interpretation of the past”. La primera como una concepción cerrada e “ingenua” del pasado, y la segunda como

concepción disciplinar. En su estudio, el autor comprobó que buena parte de los profesores eludían la parte interpretativa de la historia y se centraban en el resultado, con evidentes repercusiones en la forma en que se puede enseñar la historia.

Resultados similares se han obtenido cuando se han estudiado las concepciones del profesorado sobre la historia más reciente y controvertida. En España, Sáiz Serrano y López-Facal (2015) han indagado en el tipo de interpretaciones y narrativas de estudiantes y profesores de historia en formación sobre la historia del S.XX. En el caso de los profesores, concluyeron que, si bien hay un aumento de narrativas interpretativas críticas con respecto a los estudiantes, seguían siendo predominantes las narrativas ejemplares y cerradas de la historia.

En el contexto chileno (Toledo y Gazmuri, 2010; Toledo, Magendzo, Gutiérrez e Iglesias, 2015) se ha señalado que, si bien los profesores chilenos podían tener una concepción crítica de la historia reciente más controvertida, como relato abierto y construido a través de interpretaciones diversas, en la práctica preferían trabajar poco este tipo de temas y lo hacían mayoritariamente desde posiciones neutrales.

Recientemente Voet y De Wever (2016) han obtenido en Bélgica un resultado similar al comprobar que la mayoría de los profesores de su estudio entendían la historia como una interpretación de las evidencias históricas (*criterialism*) y muy pocos consideraban la historia como una descripción distante y objetiva (*objetivism*) o como un relato cargado de opiniones (*subjetivism*). Los investigadores concluyeron que los tres enfoques se correspondían también con diferentes formas de enseñanza de la historia, y que existía también una fuerte relación con los contextos de enseñanza.

Concepciones y creencias sobre las finalidades de enseñar historia

Por lo que respecta a las investigaciones que han puesto en relación las concepciones de la historia con las finalidades de su enseñanza, éstas han sido menos proliferas pero también interesantes. Según la investigación a través de entrevistas de McCrum (2013), la manera en que los profesores entienden la naturaleza de su disciplina interfiere en sus propósitos de enseñanza. Para este investigador, siguen predominando concepciones empiristas de la historia, y plantea dos tipos de creencias sobre la

enseñanza: unas centradas en el profesor, basadas en la transmisión del conocimiento, y otras centradas en el alumno que priorizan el razonamiento de los estudiantes.

Del mismo modo, también parece existir una fuerte relación entre la concepción de la historia y los objetivos y finalidades de su enseñanza. En esta línea, Husbands, Kitson y Pendry (2003) concluyeron que los profesores con un mayor y más profundo dominio de los contenidos eran los que mejor interpelaban a sus estudiantes y sacaban más provecho de los contenidos. Resultados similares han sido alcanzados también en contextos muy diferentes. Es el caso de Filipinas, donde Baraquiél y Oshihara (2014) comprobaron que los profesores con una visión más amplia e integradora de la historia (llamada *kasaysayang bayan*) tendían a una enseñanza con miras al desarrollo de habilidades de pensamiento histórico y de educación ciudadana.

También el estudio chileno de Toledo y Gazmuri (2010), señalado anteriormente, abordó las finalidades de enseñanza del profesorado sobre la historia reciente, y comprobó también en este caso una distancia entre los profesores y sus estudiantes. Mientras que los profesores encontraban que la historia debía servir a sus estudiantes para comprender el presente desde su historicidad, los estudiantes sólo encontraban valor a la historia por la cultura general que aporta conocer el pasado.

En definitiva, aunque el conocimiento disciplinar no sea el único de los conocimientos del profesor, la mayoría de las investigaciones relatadas muestran que su naturaleza y profundidad es determinante para la enseñanza de la historia.

La transición a la democracia en España como caso de estudio controvertido

Es habitual que la historia más cercana en el tiempo, y que representa un conflicto entre partes de una sociedad, padezca un camino incierto en el campo educativo. En los casos más graves, este tipo de historia no llega a aparecer en los currículos nacionales, o bien se muestra desde una narración mítica o impuesta, sin opción a la multiperspectividad (Paulson, 2015).

En general, se ha publicado ampliamente sobre la enseñanza de temas controvertidos y socialmente vivos de historia, su tratamiento didáctico

y el valor de introducir adecuadamente este tipo de discusiones en el aula (Barton y MacCully, 2007; Legardez y Simonneaux, 2006; Low-Beer, 1999; Psaltis, Carretero y Čehajić-Clancy, 2017). En algunos casos, se ha estudiado al profesor desde el punto de vista de sus estrategias y concepciones de enseñanza sobre temas controvertidos (Katrin, 2016; Toledo, Magendzo, Gutiérrez y Iglesias, 2015). Sin embargo, faltan investigaciones sobre cómo este tipo de contenidos históricos es procesado por el profesor, y cómo sus conocimientos y creencias pueden interactuar con las finalidades de su enseñanza. Compartimos la idea de que, cuando una sociedad se muestra dividida sobre las interpretaciones que se hacen del pasado, puede ser más importante el enfoque y las finalidades de enseñanza que el contenido en sí que se trabaje (Cole y Barsalou, 2006).

En el caso de España, la parte más controvertida de nuestra historia reciente es aquella referida a la guerra civil (1936-1939) y la dictadura franquista (1939-1975). Su tratamiento curricular y en los manuales ha ido evolucionando en paralelo a los cambios sociales y democráticos. En esta evolución, también se ha observado un cambio en las finalidades de enseñar historia, desde un origen basado en avivar los sentimientos nacionalistas (López-Facal, 2003) hasta la lógica actual de formar a los estudiantes para la interpretación del pasado y el análisis de la sociedad (Prats y Santacana, 2014). En esta transformación, el uso de metodologías activas y el enfoque por competencias, más allá de los contenidos sustantivos, han tenido un papel fundamental y es actualmente objeto de múltiples investigaciones (López-Facal, Miralles, Prats y Gómez, 2017).

Nosotros hemos elegido el caso de la transición a la democracia en España como caso controvertido por su *conexión con la actualidad y con algunos de los debates* políticos más polémicos (sistema monárquico, organización territorial, reforma de la Constitución, ley electoral, etc); por su *permanencia en la memoria* y el cuestionamiento de una política basada en el “pacto de silencio” u “olvido” (Espinosa, 2007; Kovras, 2014); y por el *enfrentamiento historiográfico* entre la corriente mayoritaria que alaba el carácter rápido y pacífico de la transición (Fusi, 2009; Linz, 1996) y los que la acusan de falta de autocrítica y de mantener un discurso oficialista y políticamente correcto (Gallego, 2008; Muniesa, 2005).

Método

Dado el carácter indagatorio y reflexivo de nuestro objetivo de investigación, se optó por un método cualitativo que favoreciera el análisis profundo del entramado de concepciones y creencias de los profesores, por encima de los resultados masivos pero más superficiales que nos podría ofrecer el uso de un método cuantitativo. Con este propósito, se decidió levantar una serie de entrevistas semiestructuradas, potenciando el desarrollo abierto de cada pregunta (Kvale, 2011).

Muestra

La muestra ha estado compuesta por un total de 39 profesores de historia de Enseñanza Secundaria y Bachillerato. El número se decidió con el propósito de obtener un cuerpo de información amplio y variado, pero que hiciera viable un estudio con suficiente profundidad de cada uno de los participantes. Éstos fueron reclutados en ciudades de siete comunidades autónomas españolas, buscando la representatividad de la centralidad (Madrid, Valladolid) y, especialmente, de la periferia (Barcelona, Bilbao, Murcia y Santiago de Compostela). Esta selección responde, por un lado, al interés de lograr la representación de diversos ámbitos geográficos y culturales a nivel nacional, y por otro, de incluir regiones en las que, por sus características históricas o por su evolución socio-política, nuestro tema de estudio pudiera ofrecer miradas y experiencias diferentes, como en los casos de Cataluña y el País Vasco. El diseño de la muestra consideró, además, las siguientes características de los informantes: a) todos los profesores impartían el periodo histórico de la transición en 4º curso de la ESO o 2º de Bachillerato; b) se consideró una proporción al 50% entre profesores mayores de 45 años, con recuerdos directos de la transición, y profesores menores de ese rango; c) se buscó una representación de profesores de centros públicos (60%) y profesores de centros privados o concertados (40%). Por lo que respecta al nivel académico de los profesores, todos contaban con el título de licenciado en Historia, pero alguno también en otras Ciencias Sociales, y bastantes poseían títulos de máster (10) o doctorado (4).

Instrumento

La condición discursiva de nuestra investigación, cuyos objetivos van encaminados a conocimientos y creencias, han determinado nuestra elección de la entrevista como el instrumento más adecuado, en la búsqueda del discurso espontáneo y libre del sujeto. De entre la amplia gama de tipología de entrevista se ha seleccionado aquella de tipo semiestructurado (Flick, 2007), que si bien mantiene el carácter abierto de la entrevista cualitativa, permite diseñar ítems previamente en forma de guion. El método de la entrevista ha sido comúnmente utilizado por otros grupos de investigación para el estudio del pensamiento del profesorado (por ejemplo Baraquiél y Oshihara, 2014; McCrum, 2013; Voet y De Weber, 2016; Yilmaz, 2010)

Un primer modelo de entrevista fue practicado con profesores de Valladolid y posteriormente sometido a grupo de expertos. La entrevista final fue aplicada a todos los profesores de la muestra, a modo de conversaciones que duraban entre 50 y 70 minutos. Las entrevistas se realizaron en ambientes íntimos y tranquilos, respetando el anonimato de los profesores para generar mayor confianza.

Las categorías que estructuraron el guion de la entrevista fueron los conocimientos y creencias sobre: (a) el contenido, conocimientos, creencias e interpretación y (d) las finalidades y propósitos de enseñanza del contenido.

Análisis de la información

Las entrevistas fueron grabadas en audio, transcritas y posteriormente volcadas al programa de análisis cualitativo Atlas.ti para facilitar su análisis con un monto total de 242.572 palabras. Siguiendo a Kvale (2011) se realizó un análisis de significados en tres fases: codificación del significado, condensación del significado e interpretación del significado. La codificación reunió aquellas ideas, conceptos, interpretaciones y proposiciones que nos servían para la consecución de nuestros objetivos de investigación. Se realizó una codificación teórica inicial (top down) y durante el transcurso del análisis emergieron nuevas categorías (bottom-up). Las categorías iniciales (ver tabla 1) surgieron de la lectura de un amplio conjunto de obras sobre el periodo de la transición a la democracia y sobre el pensamiento del profesorado de historia.

TABLA I. Codificación inicial

Familia	Código	Sub-código
Conocimientos, creencias e interpretaciones sobre el contenido	Conocimientos y creencias	Factores históricos destacados Acontecimientos destacados Conexión con el presente
	Elementos historiográficos interpretativos	Marco cronológico del periodo Violencia durante la transición Pactos y consensos
Propósitos y finalidades de enseñanza	Tipo de historia y pensamiento histórico Valores a desarrollar	

Si bien el levantamiento de las entrevistas y el análisis principal fueron realizados por el mismo investigador, un segundo investigador se encargó de la revisión paralela de las transcripciones y del análisis, con el fin de controlar la validez del estudio a través de la concordancia interpretativa.

Resultados

Para facilitar la exposición de los resultados, se ha organizado la información en función de las dos categorías iniciales principales: conocimientos, creencias e interpretaciones sobre el contenido, y conocimientos y creencias sobre las finalidades y propósitos de enseñanza.

Conocimientos, creencias e interpretaciones sobre la transición a la democracia

Para poder analizar y clasificar los conocimientos y creencias del profesorado, sus discursos históricos fueron analizados en relación a tres debates historiográficos de la transición: la violencia durante la transición, el papel de los pactos y consensos, y la cronología del periodo. En el primer caso, para una parte de la historiografía la violencia no ha sido un

elemento de estudio destacable, considerando el periodo de la transición particularmente modélico y pacífico (Fusi, 2009; Linz, 1996; Tusell, 1994). Para otros historiadores, sin embargo, la violencia fue un elemento con un peso muy importante en el proceso (Aguilar y Sánchez-Cuenca, 2009; Sánchez, 2010). Algo parecido ocurre con los pactos y consensos, cuyo carácter fue ejemplar para unos (Fusi, 2009; Linz, 1996), mientras que para otros los pactos resultaron una forma de permitir la continuidad del franquismo y de evitar la persecución de sus crímenes (Muniesa, 2005; Navarro, 2011). Finalmente, la cronología está dividida entre quienes consideran el periodo concluido en 1978 o 1982 (la mayoría de historiadores, por ejemplo Maravall y Santamaría, 1994; Soto, 2009) y aquellos que consideran que aún hoy estamos en vías de una transición real o proponen la necesidad de una segunda transición (Muniesa, 2005).

Además del análisis histórico, también se pidió a los profesores hacer una interpretación valorativa sobre el periodo de estudio, y se les preguntó si veían una conexión con la actualidad política, social o económica. Del análisis conjunto de los debates señalados y de sus interpretaciones generales se realizó una clasificación en tres grupos de pensamiento: *positivo-mítico*, *positivo-crítico* y *negativo-crítico*. Esta categorización no debe ser entendida como compartimentos estancos, sino que, al igual que en la historiografía, las reflexiones de los profesores podían mezclar ideas de los tres modelos. No obstante, en la construcción total de su discurso los profesores tendían con bastante claridad a situarse en una u otra categoría.

Una visión *positiva-mítica* de la historia

Este grupo incluye a todos aquellos profesores que consideran el periodo una etapa esencialmente modélica de la historia de España, con escasas reflexiones críticas al respecto. Suelen basarse en la idea de la transición como modelo internacional, reconocido y exportable a otras realidades, donde unos pocos artífices políticos habrían manejado magistralmente los hilos del cambio, de forma pacífica y contando con el consenso general de la población. Adolfo Suárez y el Rey aparecen como protagonistas de los pactos a los que se debe el cambio, a veces con connotaciones míticas. La cronología del periodo termina en 1978 con la aprobación de

la Constitución. De los 39 profesores entrevistados, 14 compartían esa visión de la historia.

Un ejemplo significativo es la profesora 3, quien considera el periodo como un producto histórico exportable a otras realidades, un modelo de transición del que se siente orgullosa: “una cosa bien hecha y para exportar. Es habitual que en cualquier país que nace una democracia nueva, se copia de España. Sea el papel del Rey, sea el papel de los partidos políticos, siempre se toma como referencia España” (T.3/-45 años). Nótese la carga valorativa del discurso de la profesora y cómo la concepción de la historia denota un producto concluido y cerrado, sin opción a la interpretación crítica.

Otro elemento importante en esta visión es la personificación de los procesos históricos. Los personajes políticos protagonistas llegan a adquirir connotaciones épicas y aparecen como artífices exclusivos, olvidando otros factores del análisis histórico como el contexto o la causalidad. Por ejemplo, el profesor 15 señala: “a mí la transición me parece muy bien cómo se llevó a cabo, y sobre todo cómo la llevo a cabo el Rey, que fue uno de los primeros que intentó imponer la democracia, digamos, imponer entre comillas” (T.15/+45 años).

Estas interpretaciones están cargadas de creencias, pero coinciden con el modelo presentado por algunos estudios politológicos de los años 90, que llegaron a mostrar el caso español como modelo transicional Linz, 1996). Además, coincide con la visión presentada por amplios medios de comunicación, a través de documentales y libros, como los volúmenes publicados por El Mundo o la serie televisiva de Victoria Prego, que durante buena parte de la democracia presentaron una visión del periodo muy marcada por el papel de las élites políticas y con escaso protagonismo de otros elementos históricos, como la propia ciudadanía (Quirosa-Cheyrouze, 2009).

Una visión *positiva-crítica* de la historia

Con esta categoría agrupamos aquellas posiciones que también interpretan positivamente el proceso de la Transición, pero a través de una concepción abierta y crítica de los acontecimientos. En estos casos, se hace uso de ejemplos argumentativos con los que justifican su posición a través de mecanismos de análisis más complejos que en

el grupo anterior. Por ejemplo, no destacan el carácter pacífico, pero argumentan que pudo haber sido más violento sin los pactos y consensos realizados. La cronología del periodo termina en 1978 o tras el intento fallido de golpe de Estado en 1981. Se trata del conjunto de percepciones más numeroso, agrupando a 17 profesores.

Esta concepción de la historia incluye un análisis del pasado en el que está presente el contexto y diferentes condicionantes históricos. Un condicionante histórico presentado por varios profesores de la muestra fue el terrorismo. Como apunta el profesor 27: “estando presente el tema del terrorismo de Estado y el de ETA, y el de otros grupos terroristas, la transición puso unas bases muy sólidas de democracia” (T.27/-45 años). Otro condicionante histórico repetido es el recuerdo de la guerra civil, como se aprecia en el análisis del profesor 13 cuando señala: “terminó una dictadura que había salido de una guerra, en un periodo en el que los protagonistas de la guerra aún estaban vivos, y al cabo de siete años ya había un gobierno socialista” (T.13/+45 años). Ambos condicionantes, tanto el peso de la violencia terrorista como el recuerdo de la guerra civil, son argumentos ampliamente repetidos desde la argumentación historiográfica más moderada (Maravall y Santamaría, 1994).

Una visión negativa-crítica de la historia

Una tercera opción agrupa al profesorado abiertamente crítico tanto del proceso histórico como de su resultado. Expresan un razonamiento histórico abierto e interpretativo, como el caso anterior, pero sus argumentaciones y reflexiones transcurren por caminos muy diferentes. Suelen centrar su discurso en temas conflictivos, como la no persecución de los crímenes franquistas o las continuidades políticas y económicas del régimen. La interpretación crítica se realiza también en este caso desde la contextualización de los acontecimientos, pero considerando otros elementos interpretativos. El modelo de transición pactada aparece ahora desde un sentido negativo, presente también en la historiografía (Espinosa, 2007). Además, coincidiendo con la historiografía más crítica, la violencia cobra mayor importancia en sus discursos y, en algunos casos, el periodo de transición no ha concluido. Hasta 8 de los docentes entrevistados participarían de este enfoque.

En relación a los pactos, por ejemplo, el profesor 34 argumenta que, habiendo nacido desde el interior de la propia dictadura franquista, “el franquismo legitimó la democracia, y la democracia legitimó el franquismo” (T.34/+45). Los pactos implicaron, para estos profesores, el fin de las reivindicaciones de la izquierda política, que cedieron sus aspiraciones para participar del proceso. Es el caso del profesor 5, con una reflexión cargada de desencanto: “la movilización ciudadana y aquellas organizaciones que realmente realizaron una lucha antifranquista, ¿de qué sirvieron? De nada, de muy poco” (T.5/-45). Otra consecuencia de los pactos habría sido la continuidad de elementos y personajes de la dictadura durante la democracia. Por ejemplo, el profesor 33 señala: “hubo continuidad de ministros, de gente que estaba en primera línea, en la Guardia Civil... que continuaron en democracia sin tener que dar ninguna explicación” (T.33/-45 años). Un argumento que ha sido corroborado en el ámbito historiográfico y de la investigación periodística (Grimaldos, 2004) y que conectaría con la carencia de un proceso de justicia transicional (Kovras, 2014).

Es importante señalar que la gran mayoría de las visiones negativas-críticas se detectaron en las comunidades históricas (Galicia, País Vasco y Cataluña), con movimientos políticos independentistas. Una postura que, en muchos casos, está unida únicamente a discursos marcadamente críticos con el proceso de transición a nivel nacional, descuidando la reflexión interna sobre el reparto de poder entre partidos independentistas en estas comunidades, resultante también del mismo proceso de transición. El factor ideológico y territorial se presenta así como una variable a tener en cuenta en la configuración del pensamiento del profesorado sobre este periodo.

Finalidades y propósitos de enseñanza y su relación con los conocimientos, creencias e interpretaciones de la historia

A lo largo de las entrevistas, los profesores fueron interrogados sobre los propósitos y finalidades que encontraban en la enseñanza de la historia y, particularmente, en la enseñanza del periodo de la transición. Por lo general, los profesores de la muestra encuentran un fuerte valor en la enseñanza de la historia como medio para que sus estudiantes comprendan mejor el presente y actúen para cambiarlo. Sin embargo,

cuando profundizamos en esta visión, encontramos una importante distancia entre el discurso ideal y los ejemplos concretos con los que argumentan su posición. Es decir, incluso muchos de los profesores con una concepción de la historia crítica y conectada con el presente, encontraban el valor de su enseñanza en un relato ejemplar y cerrado del pasado, pocas veces interpretativo o crítico.

TABLA II. Concepciones de la historia y finalidades de enseñanza

Concepción de la historia	Conocimientos, creencias e interpretaciones	Historia a enseñar/ propósito/valores	Nº de profesores
Positiva-mítica	Relato épico de pactos y personajes históricos. Transición modélica. Transición pacífica. Fecha de fin: 1978.	<i>Historia ejemplar.</i> Comprensión del presente. Valores: consenso, unidad nacional y constitución, paz, democracia.	14
Positiva-crítica	Interpretación crítica. Condicionantes históricos: violencia terrorista, recuerdo de la guerra civil. Fecha de fin: 1978/1981.	<i>Historia ejemplar/historia problema.</i> Comprensión del presente. Valores: consenso, paz, participación, democracia.	17
Negativa-crítica	Interpretación crítica. Condicionantes históricos: falta de justicia transicional, cesión de la izquierda. Fecha de fin: 1981/no ha terminado.	<i>Historia ejemplar/historia problema.</i> Comprensión del presente. Valores: cesiones, justicia transicional, democracia.	8

Como se puede observar en la Tabla 2, tanto el tipo de interpretaciones, como las creencias y valores que los profesores asocian al periodo histórico, no siempre tienen un reflejo en las finalidades de enseñanza. Podemos dividir las finalidades en dos grandes grupos: en relación al tipo de pensamiento histórico que se pretende desarrollar (*historia ejemplar e historia problema*) y en relación a los valores asociados a su enseñanza.

Historia ejemplar o historia problema

Cuando analizamos los discursos de los profesores participantes surgen dos formas de comprender la finalidad de enseñar historia, que corresponden a dos formas de pensar la historia: la primera y más común, que hemos denominado *historia ejemplar*, entiende la historia como relato que contiene en sí mismo una enseñanza; la segunda, que llamamos *historia problema*, encuentra su valor formativo no en el relato, sino en el análisis o interpretación que se haga del pasado.

Como se puede apreciar en la tabla 2, todos los profesores con una concepción *positiva-mítica* de la historia desarrollaron discursos sobre las finalidades de enseñanza fácilmente clasificables dentro de la *historia ejemplar*. Por ejemplo, si seleccionamos el profesor 15, observamos que la falta de razonamiento histórico complejo coincide con una concepción simplista y ejemplar de lo que espera enseñar: “si ellos [los alumnos] estudian el comienzo de la democracia y las normas y las leyes dadas por la Constitución, es como darles unas pautas para que se comporten de una forma correcta y sean buenos ciudadanos” (T.15/+45).

Sin embargo, sorprende comprobar que una buena parte del profesorado con concepciones *positiva-crítica* y *negativa-crítica*, también encuentran sus finalidades de enseñanza en la *historia ejemplar*. En estos casos, existe una incongruencia entre su forma de entender la historia, de carácter crítico e interpretativo, y sus finalidades de enseñarla, como relato ejemplar. Esta incoherencia se justifica, en algunos casos, porque los profesores consideran que la historia de la transición, por su carácter complejo y reciente, no puede ser trabajada en profundidad con los alumnos y que debe plantearse en términos muy concretos. Así, el profesor 21 dice: “hay que explicarles la diferencia entre una dictadura y una democracia, los valores democráticos. Pero no se puede trabajar en profundidad porque puedes confundir a los alumnos, y hay que tener mucho cuidado” (T.21/+45). Es decir, la explicación de un relato cerrado y simplificado evita a los profesores tener que entrar en debates sobre el pasado reciente, suprimiendo así también buena parte de su valor formativo.

Son aquellos docentes con una visión más crítica del pasado transicional (sea *positiva-crítica* o *negativa-crítica*) los que demuestran una mayor complejidad en sus finalidades de aprendizaje, que podemos definir como *historia problema*. En estos casos, sus finalidades de enseñanza

conectan con su concepción interpretativa y crítica de la historia. Por ejemplo, el profesor 34 señala que es importante enseñar “a adoptar un punto de vista crítico, a entender las luces y las sobras de lo que fue pactar, negociar, renunciar...” (T.34/-45). Además, la conexión crítica que hacían estos profesores con el presente al interpretar la historia, también aparece dentro de sus finalidades de enseñanza. Los estudiantes, señala el profesor 5, “deben entender en qué régimen político y legal viven ahora; y no sólo la Constitución como tal, sino cómo fue hecha la Constitución, en qué sentido fue redactada y fue discutida” (T.5/+45).

Valores asociados al aprendizaje de la historia

Otro elemento importante del análisis fueron los valores que los profesores del estudio declaraban enseñar como parte de sus finalidades. Como también se puede observar en la tabla 2, los profesores con una concepción *positiva-mítica* de la historia destacaron valores como el consenso de la clase política, la unidad nacional, la resolución pacífica de los conflictos y el triunfo de la democracia. Por ejemplo, el profesor 2 dice que el estudio de la transición sirve para enseñar “el consenso, una nueva organización del Estado, el orden de la Constitución y el valor de la clase política comprometida con el cambio y con la mejora del país” (T.2/-45).

En el caso de los profesores con una concepción *positiva-crítica* de la historia, comparten algunos valores a enseñar con el grupo anterior, como es el consenso, la resolución pacífica de los conflictos y el triunfo de la democracia. Pero aparece un valor nuevo: la participación, confiriendo importancia al papel de la ciudadanía y no sólo a la clase política. Es el caso del profesor 27 cuando dice: “ahora estamos un poquito más aborregados. No siendo modélica, creo que [la transición] fue un momento de participación, de implicación, de lucha, de ser críticos como ciudadanos” (T.27/+45).

Los valores positivos aparecen incluso en los profesores con una visión *negativa-crítica* de la historia. Ciertamente, no se comparten valores como la unidad nacional o el consenso, pero se rescató el valor de la democracia sobre la dictadura y se limitó el discurso crítico a la hora de hablar de finalidades de enseñanza. Algunos profesores expresaron un valor crítico-negativo en la enseñanza del periodo: el valor de la justicia,

que para el caso de la transición se explica como un valor fallido. Así, el profesor 7 dice: “el aprendizaje es que si hay una transición lo mejor es hacerlo de forma pacífica, pero siempre con justicia. Lo que no puede ser es que los propios verdugos sigan en el poder. A mí me parece injusto” (T.7/+45).

En definitiva, valores, creencias e interpretaciones siguen una argumentación que se traspaesa desde las concepciones de la historia hasta las finalidades de su enseñanza.

Discusión

Nuestra clasificación del pensamiento docente en tres categorías tiene amplias similitudes con la distinción que Yilmaz (2010) hace entre “history as the past” and “history as an interpretation of the past”. En nuestro caso, sin embargo, la mayoría de los profesores entrevistados abordaron la historia haciendo una interpretación del pasado (17 de concepción *positiva-crítica* y 8 de concepción *negativa-crítica*), aunque existe un número no pequeño de profesores con una concepción cerrada de la historia, como simple relato del pasado (14 de concepción *positiva-mítica*). Estos resultados coincidiría con el trabajo de Voet y De Wever (2016) si observamos que nuestra clasificación *positiva-crítica* y *negativa-crítica* puede ser identificada con la opción criterialista de ese estudio.

Respecto al estudio cuantitativo realizado en España por Sáiz Serrano y López-Facal (2015), existen similitudes con nuestros resultados por cuanto buena parte de los profesores expresan relatos de la historia suficientemente ricos, no lineales y con interrelación de ámbitos de análisis (social, político, económico...). Pero, al igual que en el estudio de estos investigadores, buena parte del profesorado entrevistado no desarrolló una visión interpretativa crítica de aspectos controvertidos del pasado reciente.

Los estudios realizados en Chile, sobre un contexto de historia reciente similar al español, muestran un profesor de historia más crítico que el de nuestra investigación (Toledo y Gazmuri, 2010). Sin embargo, coincidimos con el estudio de Toledo y Gazmuri en que los profesores expresaron mayoritariamente que la finalidad de enseñar historia estaba en que sus estudiantes comprendieran mejor el presente. Un modelo de profesor que coincide con los estudios realizados por Evans (1990),

y que ya entonces destacaban la figura del profesor *relativist/reformer*, preocupado por generar un cambio social con su trabajo.

Estamos de acuerdo con Toledo, Magendzo, Gutiérrez e Iglesias (2015) en que el profesorado tiene unas altas expectativas a la hora de definir sus finalidades de enseñanza de la historia reciente, que no se corresponden con los medios declarados para lograrlas. De manera que muchos de los profesores entrevistados, incluso con una concepción interpretativa y crítica de la historia, valoraban el aprendizaje de la historia como relato ejemplar del pasado, no como interpretación o problematización que pudiera servir a la comprensión de presente. Es decir, se produce un desfase entre la aspiración ideal del profesor y el medio para alcanzarla.

Por lo que respecta a los valores asociados a la enseñanza de la historia, éstos fueron una parte importante de las finalidades educativas de los profesores de historia de la muestra. Esto pudo deberse a las propias características del periodo histórico, puesto que la transición representa el comienzo de la democracia después de una larga dictadura. De manera que, aunque muchos profesores realizaron una interpretación crítica, incluso negativa, del periodo histórico, a la hora de exponer sus finalidades casi todos trataron de rescatar valores positivos para su enseñanza. Sobresale así el rol del docente educador (centrado en hábitos de conducta y valores), frente al rol del experto o especialista (dedicado a la enseñanza del saber disciplinar), siguiendo la clasificación de Monereo y Monte (2011).

Coincidimos con McCrum (2013) y con Baraquiel y Oshihara (2014) en que existe una relación entre las concepciones de la historia y las finalidades de enseñanza, pero con matices. A diferencia de la investigación de Baraquiel y Oshihara (2014), donde los profesores con una concepción más amplia y crítica del pasado tenían una visión de su enseñanza más rica y compleja (con miras al desarrollo de habilidades de pensamiento histórico), en nuestro caso no siempre fue así. Muchos de los profesores participantes tendieron a limitar su concepción crítica e interpretativa de la historia al definir las finalidades de su enseñanza, reduciendo la carga analítica sobre el pasado y basándose más en un relato cerrado.

Conclusiones

Sobre los conocimientos, creencias e interpretaciones de la historia. El profesorado entrevistado demostró conocimientos muy generales sobre el pasado reciente que reforzaban con creencias y valoraciones, presentes en el debate social y mediático actual. Sus conocimientos, creencias e interpretaciones sobre la historia de la transición se correspondieron con los debates historiográficos presentados al comienzo. En base a esto, y al tipo de razonamiento histórico que los profesores profesaban, se clasificó al profesorado participante en tres grupos: *positivo-mítico*, *positivo-crítico* y *negativo-crítico*.

En total, una mayoría de los profesores expresaron algún tipo de interpretación crítica (17 de opción *positiva-crítica* y 8 de opción *negativa-crítica*), desarrollando un pensamiento complejo de argumentación y análisis histórico. Sin embargo, no es menor el hecho de que hasta 14 profesores (de opción *positiva-mítica*) asumieron un relato histórico cerrado, sin desarrollar mayor argumentación interpretativa.

La variable edad no arrojó resultados destacables, existiendo profesores jóvenes y mayores de 45 años en cada una de las categorías. Sin embargo, sí se observaron resultados respecto a la variable territorial. Predominando la visión *negativa-crítica* en regiones donde se desarrollan movimientos regionalistas o independentistas, muy críticos con el modelo de transición a la democracia; mientras que las visiones positivas fueron mayoritarias en regiones donde no existe esta problemática. Este aspecto demuestra el peso de las corrientes de pensamiento actuales y de los medios de comunicación sobre el tipo de interpretación histórica y las creencias de los profesores entrevistados.

Sobre la relación entre conocimientos, creencias e interpretaciones y las finalidades y propósitos de enseñanza. En el caso del profesorado con una visión *positiva-mítica*, se observa una relación clara entre su visión y las finalidades de enseñanza. El mismo relato cerrado y los mismos valores de su interpretación histórica se repiten en sus discursos sobre las finalidades. Prima la transmisión de una historia ejemplar, donde el propio relato ofrece en sí el aprendizaje, unido a una serie de valores positivos como el consenso, la paz, la unidad nacional y la democracia.

Más complejo fue el caso de los profesores con visiones críticas, positivas o negativas, del pasado. Aunque se dio una relación entre visiones críticas y finalidades más profundas y complejas de la enseñanza

de la historia, en muchos casos, a la hora de establecer las finalidades educativas de enseñar la historia, se perdió el discurso crítico y el trabajo de interpretación más complejo. En definitiva, la historia enseñada parece tender a simplificarse y a convertirse en un relato ejemplar, que asume valores homogéneos y, a veces, superficiales; limitándose una finalidad principal de la enseñanza de la historia, como es el desarrollo de habilidades de pensamiento histórico y pensamiento crítico a través del análisis multicausal y desde múltiples perspectivas del pasado (Prats y Santacana, 2014; Rösen, 2005; Sáiz y López-Facal, 2015). Sólo algunos profesores llegaron a plantear el foco de su finalidad educativa, no en el relato histórico, sino en la interpretación y problematización del pasado. Estos casos estuvieron más presentes dentro de la concepción *negativa-crítica*, lo que nos lleva a considerar que las valoraciones del pasado interfieren en las finalidades de su enseñanza.

Como proyección del estudio, consideramos importante ampliar la investigación con un estudio de aula. Es decir, analizar lo que finalmente ocurre en las salas de clase cuando se trabaja esta temática cercana en el tiempo y de carácter controvertido, cómo abordan los profesores estos temas y si son congruentes con sus concepciones, creencias y finalidades. Por otro lado, la investigación nos muestra un profesorado con problemas para convertir sus propias habilidades históricas en aprendizajes esperados. De manera que futuras líneas de investigación deberían centrarse también en las estrategias y materiales que sirvan al profesor para el desarrollo de habilidades de pensamiento histórico.

Referencias bibliográficas

- Aguilar, P. y Sánchez-Cuenca, I. (2009). Violencia política y movilización social. En Baby, S., Compagnon, O., y González, E. *Violencia y transiciones políticas a finales del siglo XX* (pp. 95-111). Madrid: Casa de Velázquez.
- Baraquiel A. C. y Oshihara, S. (2014). Teachers' Pedagogy and Conceptions of History: Decolonizing and Transforming History in Elementary. *The SIJ Transactions on Industrial, Financial & Business Management (IFBM)*, 2(3), 87-94.

- Barton, K. C. y McCully, A. W. (2007). Teaching controversial issues... where controversial issues really matter' in Teaching History. *Teaching History. The Historical Association*, 127, 13-19.
- Bohan C. y Davis O. (1998). How social studies student teachers historical thinking is reflected in their writing of history. *Theory and research in social education*, 26, 173-197.
- Buehl, M. M., y Beck, J. (2015). The relationship between teachers' beliefs and practices. En H. Fives y M. Gregorie Gill (Eds.), *International handbook of research on teachers' beliefs* (pp. 66-84). New York, NY: Routledge
- Carretero, M., Berger, S. y Grever, M. (2017). Introduction: Historical Cultures and Education in Transition. En Carretero, Berger y Grever (Eds.). *Palgrave Handbook of Research in Historical Culture and Education* (pp. 1-37). Palgrave Macmillan.
- Cole, E., y Barsalou, J. (2006). *Unite or Divide? The Challenges of Teaching History in Societies Emerging from Violent Conflict. Special report*. Washington, DC: U.S. Institute of Peace.
- Espinosa, F. (2007). De saturaciones y olvidos. Reflexiones en torno a un pasado que no puede pasar. *Hispania Nova. Revista de Historia Contemporánea*, 7, 413-440.
- Evans, R. W. (1990). Teachers' conceptions of history revisited: Ideology, curriculum and student belief. *Theory and Research in Social Education*, 18(2), 101-138.
- Fives, H., & Buehl, M. M. (2016). Teachers' Beliefs, in the Context of Policy Reform. *Policy Insights from the Behavioral and Brain Sciences*, 3(1), 114-121.
- Flick, U. (2007). *Introducción a la investigación cualitativa*. Madrid: Morata.
- Fusi, J. P. (2009). España 1975-2008. In Carr, R. *España 1808-2008* (epilogue). Barcelona: Ariel.
- Gallego, F. (2008). *El mito de la transición. La crisis del franquismo y los orígenes de la democracia (1973-1977)*. Barcelona: Crítica.
- Gimeno, J., y Pérez, A. (1988) Pensamiento y acción en el profesor: de los estudios sobre la planificación al pensamiento práctico. *Infancia y Aprendizaje*, 42, 37-63.
- Grimaldos, A. (2004). *La sombra de Franco en la transición*. Madrid: Oberon.

- Grossman, P. L. (1990). *The making of a teacher: Teacher knowledge and teacher education*. New York: Teachers College Press.
- Husbands, C., Kitson, A. y Pendry, A. (2003) *Understanding history teaching: teaching and learning about the past in secondary schools*, Maidenhead: Open University Press.
- Kagan, D. M. (1992). Implications of research on teacher belief. *Educational Psychologist*, 27(1), 65-90.
- Katrin, K. (2016). Sensitive and controversial issues in the classroom: teaching history in a divided society. *Teachers and Teaching*, 22(1), 35-53.
- Kvale, S. (2011). *Las entrevistas en investigación cualitativa*. Madrid: Morata.
- Kovras, I. (2014). *Truth Recovery and Transitional Justice. Deferring human rights issues*. New York: Routledge.
- Legardez, A. y Simonneaux A. (2006). *L'école à l'épreuve de l'actualité. Enseigner les questions vives*. Issy les Moulineaux: ESF éditeur,
- Linz, J. (1996). La transición española en perspectiva comparada. En Tusell, J. y Soto, A (eds.). *Historia de la transición. 1975-1986* (pp. 21-45). Madrid: Alianza Editorial.
- López-Facal, R. (2003). *La enseñanza de la historia, más allá del nacionalismo*. Madrid: Marcial Pons Historia.
- López-Facal, R., Miralles, P., Prats, J. (Dir.) y Gómez, C. (Coord.) (2017). *Enseñanza de la historia y competencias educativas*. Barcelona: Graó.
- Low-Ber, A. (1999). *Teaching controversial and sensitive issues in history education for secondary schools*. Strasbourg: Council of Europe.
- Maravall, J. M. y Santamaría, J. (1994). El cambio político en España y las perspectivas de la democracia. En O'Donnell, G., Schmitter, P. & Whitehead, L. *Transiciones desde un gobierno autoritario, 1. Europa meridional* (pp. 112-164). Barcelona: Paidós.
- McCrum, E. (2013). History teachers' thinking about the nature of their subject. *Teaching and Teacher Education*, 35(1), 73-80.
- Monereo, C. y Monte, M. (2011). *Docentes en tránsito. Incidentes críticos en secundaria*. Barcelona: Graó.
- Muniesa, B. (2005). *Dictadura y Transición. La España lampedusiana. II: La monarquía parlamentaria*. Barcelona: Universitat de Barcelona.
- Ortiz-Heras, M. (2004) Historiografía de la transición. En *La transición a la democracia en España. Historia y fuentes documentales. VI*

- Jornadas de Castilla La Mancha sobre investigación en archivos* (pp. 223-240). Guadalajara: Anabad.
- Paulson, J. (2015). "Whether and how?" History education about recent and ongoing conflict: a review of research. *Journal on Education in Emergencies*, 1(1), 14-47.
- Prats, J. y Santacana, J. (2014). Por qué y para qué enseñar Historia. En Prats, J. (coord.). *Didáctica de la Geografía y la Historia* (pp. 13-27). Barcelona: Graó.
- Psaltis, C., Carretero, M. y Čehajić-Clanc, S. (2017) (Eds.). *History Education and Conflict Transformation*. Palgrave Macmillan.
- Sáiz-Serrano, J. y López-Facal, R. (2015). Competencias y narrativas históricas: el pensamiento histórico de estudiantes y futuros profesores españoles de educación secundaria. *Revista de Estudios Sociales*, 52, 87-101.
- Quirosa-Cheyrouze y Muñoz, R. (2009). Los medios de comunicación ante el proceso democratizador en España. En Quirosa-Cheyrouze y Muñoz, R. (Ed.). *Prensa y democracia. Los medios de comunicación en la Transición* (pp. 13-25). Madrid: Biblioteca Nueva.
- Rüsen, J. (2005). *History: Narration, Interpretation, Orientation*. Nueva York: Berghahn.
- Sánchez, M. (2010). *La transición sangrienta. Una historia violenta del proceso democrático en España (1975-1983)*. Barcelona: Grup Editorial.
- Shulman L.S. (1987). Knowledge and teaching: Foundatios of the new reform. *Harvard educational review*, 57(1), 1-21.
- Soto, A. (2009). Violencia política y transiciones a la democracia. En Baby, S., Compagnon, O.y& González Calleja, E. (Eds.) *Violencia y transiciones políticas a finales del siglo XX* (pp. 113-129). Madrid: Casa de Velázquez.
- Toledo, M.I. y Gazmuri, R. (2010). Sobre la enseñanza y el aprendizaje de la dictadura militar y la transición a la democracia en las aulas de secundaria en Santiago de Chile. En *VII Jornades Internacionals de Recerca en Didàctica de les Ciències Socials*. Universidad Autónoma de Barcelona.
- Toledo, M. I., Magendzo, A., Gutiérrez, V. y Iglesias, R. (2015). Enseñanza de 'temas controversiales' en la asignatura de historia y ciencias sociales desde la perspectiva de los profesores. *Estudios pedagógicos*, 41(1), 275-292.

- Tusell, J. (1994). La transición a la democracia en España como fenómeno de Historia política. *Ayer. Revista de Historia Contemporánea*, 15, 55-76.
- Virta, A. (2001). Student teachers' conceptions of history. *International Journal of Historical Learning, Teaching and Research*, 2(1), 53-64.
- Voet, M. & De Wever, B. (2016). History teachers' conceptions of inquiry-based learning, beliefs about the nature of history, and their relation to the classroom context. *Teaching and Teacher Education*, 55(1), 57-67.
- Wineburg, S. (1991) Historical problem solving. A study of the cognitive processes used in the evaluation of documentary and pictorial evidence. *Journal of Educational Psychology*, 83, 73-87.
- Yeager E. y Davis O. (1996). Classroom teachers' thinking about historical texts: an exploratory study. *Theory and research in social education*, 24(2), 146-166.
- Yilmaz, K. (2010). Social studies teachers' conceptions of history: Calling on historiography. *Journal of Educational Research*, 101(3), 37-41.
- Dirección de contacto: Rosendo Martínez Rodríguez. Universidad de Valladolid. Facultad de Educación de Segovia. Plaza de la Universidad, 1. 40005, Segovia. E-mail: rosendo.martinez@uva.es.

Pensémoslo de nuevo: ¿Podemos comparar las escalas de antecedentes socioeconómicos?!

Back to the drawing board: Can we compare socioeconomic background scales?

DOI: 10.4438/1988-592X-RE-2019-383-400

Andrés Sandoval-Hernandez

University of Bath

David Rutkowski

Indiana University

University of Oslo

Tyler Matta

Pearson

University of Oslo

Daniel Miranda

Pontificia Universidad Católica de Chile, MIDE UC

Centro de Estudios de Conflicto y Cohesión Social

Resumen

Utilizando datos de evaluaciones internacionales de gran escala, evaluamos la consistencia y la invarianza de las escalas de antecedentes socioeconómicos de los estudiantes entre los países participantes en estos estudios. Para ello, utilizamos las medidas de antecedentes socioeconómicos desarrolladas por PISA, TERCE y TIMSS, ya que cada estudio operacionaliza esta medida de manera diferente. Como parte de nuestro análisis, examinamos si la escala de TERCE, un estudio latinoamericano –con medidas que fueron desarrolladas con

⁽¹⁾ Este artículo fue apoyado por el Norwegian Research Council Grant #255246 y por el Centro de Estudios de Conflicto y Cohesión Social —COES CONICYT/FONDAP N°15130009

enfoque regional– exhibe mejores propiedades psicométricas que otras medidas que fueron diseñadas para funcionar en un número mayor y más diverso de sistemas educativos. Adicionalmente, examinamos la escala de TIMSS, un estudio enfocado en tendencias –que históricamente ha enfatizado la comparabilidad y consistencia–. Finalmente, incluimos también la escala de PISA, que cuenta con el mayor número de participantes y que, en cierta medida, ha cambiado y conceptualizado sus cuestionarios de contexto dependiendo el dominio principal y el foco de cada ciclo del estudio. Nuestros resultados sugieren que ninguna de las escalas de contexto que analizamos son completamente invariantes entre los países que participen en cada estudio, y por lo tanto las comparaciones entre países deben hacerse con precaución. Este artículo discute los niveles de equivalencia alcanzados por cada escala en cada estudio, así como el tipo de comparaciones que se pueden realizar dados estos resultados (e.g. comparación de los promedios nacionales de las escalas, comparación de relaciones o correlaciones entre las escalas evaluadas y otras variables, etc.).

Palabras clave: medición de la invarianza, medición de la equivalencia, TERCE, TIMSS, PISA, análisis factorial confirmatorio multigrupo, escalas de antecedentes socioeconómicos.

Abstract

Using data from international large-scale assessments (ILSA), we evaluate the issue of country-level model-data consistency of background socio-economic scales, as well as the invariance across countries. To that end, we use data from PISA, TERCE, and TIMSS, as they operationalize socio-economic status somewhat differently. As part of our analysis, we examine whether TERCE, a Latin American study – with measures that are regionally developed – exhibits better psychometric properties than measures that are designed to function across a larger and more diverse number of educational systems. We also examine TIMSS, a trends focused study – that has historically emphasized consistency and comparison. Finally, we include PISA which has the largest number of participants and has changed and conceptualized a great deal of its background questionnaire depending on the study's major domain and focus. Our findings suggest that none of the socio-economic background scales we analyzed are fully invariant in any of the three studies, and therefore comparisons across countries should be done with caution. The different levels of equivalence reached by each scale in each study and the type of comparisons that can be made given these results (e.g., comparison of average scale scores, comparison of relationships between the tested scales and other variables) are discussed in the full paper.

Key words: measurement invariance, measurement equivalence, TERCE, TIMSS, PISA, multi-group confirmatory factor analysis, socio-economic scales.

Introducción

Las evaluaciones internacionales de gran escala (ILSAs, por sus siglas en inglés) acerca del logro educativo, como el Programa para la Evaluación Internacional de Alumnos (PISA), el Estudio Internacional de Tendencias en Matemática y Ciencias (TIMSS) y el Tercer Estudio Regional Comparativo y Explicativo (TERCE), tienen múltiples propósitos. Desde monitorear sistemas educativos y realizar análisis comparativos, hasta proveer de información sobre lo que los estudiantes saben y pueden hacer. Los ILSAs ofrecen a las partes interesadas una oportunidad de entender el contexto y los factores asociados al aprendizaje, y entregan antecedentes acerca de los estudiantes, los profesores y las escuelas. No obstante, a medida que aumenta la participación en estas evaluaciones, se vuelve más difícil para las organizaciones adaptarlas y cumplir con las necesidades de un conjunto cada vez más heterogéneo de países. Por ejemplo, en la evaluación PISA 2015 participaron 62 países, 34 de ellos corresponden a los países miembros de la OCDE (que representan las mayores economías en el mundo), mientras que los 28 participantes restantes (denominados países asociados) correspondieron a un conjunto heterogéneo de economías y culturas, incluyendo sistemas educacionales como Túnez, Perú, Singapur y Shanghai, China. TIMSS enfrenta una situación similar. Finalmente, aunque las evaluaciones regionales como TERCE cuentan con una menor cantidad de participantes y –ostensiblemente– presentan menor heterogeneidad que las evaluaciones internacionales más globales, la diversidad de idiomas, economías y culturas persiste tanto entre los países participantes como dentro de ellos. Por ejemplo, el PIB per cápita de Chile triplica el de Bolivia y, aunque la mayoría de los países comparten el idioma español, en muchos de los países conviven distintos pueblos indígenas con diferentes lenguas nativas.

La mayoría de los ILSAs incluyen tanto una evaluación cognitiva como un conjunto de cuestionarios de contexto. Los cuestionarios se administran a estudiantes y, dependiendo de la evaluación, pueden incorporar a profesores, padres y directores de escuelas. En general, los cuestionarios de contexto tienen dos usos principales: (1) ayudar a contextualizar el sistema educacional evaluado; y (2) optimizar la estimación del logro académico en la población y un conjunto de subpoblaciones. Los beneficios de usar datos de contexto para mejorar la estimación del logro académico han sido ampliamente documentados

(Mislevy, Beaton, Kaplan y Sheehan, 1992) y no constituyen el foco de este artículo. De hecho, algunos investigadores han destacado potenciales desafíos metodológicos asociados al amalgamamiento de participantes, apuntando especialmente al modelo de estimación del logro y a si las comparaciones son razonables y válidas cuando los sistemas educativos difieren sustancialmente (Goldstein, 2004; Kreiner y Christensen, 2014; Mazzeo y von Davier, 2009; Oliveri y Ercikan, 2011). En parte como respuesta a estas y otras críticas, el proyecto PISA ha implementado ajustes, especialmente orientados hacia participantes de bajo rendimiento (e.g., incorporando ítems de menor dificultad a los instrumentos en países donde de espera bajo desempeño; OCDE, 2012). Estudios recientes han demostrado que estos tipos de ajustes son una manera promisorio de dar cuenta y enfrentar la heterogeneidad que está inevitablemente presente en la investigación comparativa (Rutkowski, Rutkowski y Zhou, 2016). Se ha realizado un gran esfuerzo para evaluar y asegurar la comparabilidad de las escalas de logro entre países (e.g. OECD, 2014; Schulz, Ainley y Frailon, 2011; UNESCO-OREALC, 2016) y a lo largo del tiempo (e.g. Gaviria y Covadonga, 2007). Por el contrario, se ha puesto mucho menor esfuerzo en diseñar escalas derivadas de los cuestionarios de contexto que consideren diferencias importantes entre los participantes (Rutkowski y Rutkowski, 2010).

Diversas investigaciones han demostrado empíricamente que el supuesto de equivalencia entre las escalas de contexto de los ILSAs a menudo no se sostiene, lo que compromete la comparabilidad (Caro, Sandoval-Hernandez y Lüdtke, 2016; Glas y Jehangir, 2014; Oliveri y von Davier, 2014). De esta manera, el objetivo de este artículo es doble: primero, mostrar un método para explorar tanto la consistencia de las escalas de contexto dentro de los países, como la equivalencia de estas escalas entre los países. El segundo objetivo es discutir los resultados de la aplicación de este método en el caso de las escalas de nivel socioeconómico de PISA, TIMSS y TERCE. Específicamente, exploramos los diferentes niveles de comparabilidad alcanzados por las escalas utilizadas en cada estudio para medir alguna forma de nivel socioeconómico (NSE) y discutimos el tipo de comparaciones que se pueden realizar en virtud de estos resultados (e.g., comparación de las medias de las puntuaciones, comparación de las relaciones entre las escalas evaluadas y otras variables).

Dado que no sería factible evaluar la comparabilidad de todas las escalas de contexto de los tres estudios, en este artículo nos enfocamos

en las escalas desarrolladas por las organizaciones implementadoras para examinar alguna forma de NSE en tres estudios internacionales (PISA, TERCE y TIMSS). Decidimos usar estas escalas porque, en los estudios enfocados en identificar factores asociados a los resultados de aprendizaje (e.g., efectividad escolar y docente), NSE es la variable control que consistentemente muestra asociaciones más fuertes con el logro educativo. Además, existe un cuerpo importante de literatura focalizada especialmente en entender los mecanismos mediante los cuales los antecedentes socioeconómicos o el estatus socioeconómico de la familia están asociados al logro educativo (Buchmann, 2002).

Al examinar la equivalencia de estas escalas entre los países y al comparar los resultados de los diferentes estudios, podemos determinar si un diseño o enfoque de evaluación diferente resulta en distintos grados de comparabilidad. Estos tres estudios fueron elegidos intencionalmente ya que representan tres diseños de evaluación internacional diferentes. TERCE se escogió por representar un estudio regional, con preguntas que son desarrolladas a nivel regional y bajo el supuesto de que los desarrolladores del instrumento construyeron la escala para un grupo menor de participantes (Treviño, Fraser, Meyer, Morawietz, Inostroza y Naranjo, 2015). TIMSS se escogió por representar un estudio enfocado en tendencias – que ha históricamente puesto énfasis en la consistencia y la comparación de los cambios en las sociedades, constructos o participantes. Finalmente, incluimos PISA, que tiene el mayor número de participantes e históricamente ha demostrado interés en realizar cambios significativos a sus cuestionarios de contexto (OECD, 2016a).

Marco analítico

Nuestro marco analítico se sitúa a nivel general en la teoría de la medición y específicamente en la teoría y diseño de instrumentos de medición (e.g., van der Linden, 2005; Wilson, 2005). La teoría de los tests se focaliza en cómo un conjunto de respuestas observadas da cuenta de un constructo teórico, no observable. Dentro de los ILSAs, estas respuestas observables se obtienen a partir una prueba o instrumento (estandarizado), el que puede ser definido como “una técnica para relacionar algo que podemos observar en el mundo real (a veces denominado manifiesto u observable) con algo que medimos y que sólo existe como parte de una teoría (a

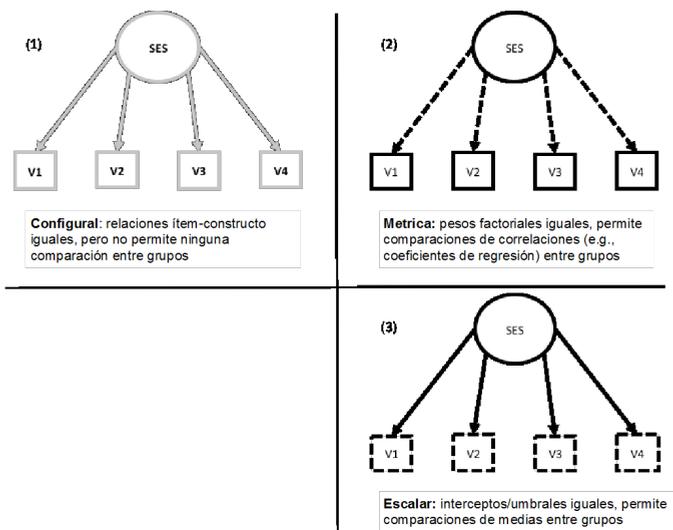
veces denominado latente o no observable)” (Wilson, p. 4). El diseño de instrumentos, o el proceso de desarrollar ítems que eliciten un constructo teórico no-observable, es un proceso iterativo. El supuesto que subyace a la teoría de los tests, y que también rige el diseño de los instrumentos, es que la relación entre los constructos teóricos y las respuestas observables a los ítems del instrumento es de tipo causal (Wilson, 2005). Esto es, el grado en que el participante presenta determinado constructo (o constructos en el caso multidimensional) provoca las respuestas a un conjunto de ítems. Debido a que no podemos observar el constructo directamente, el agente causal es latente y la medida (ítem) está construida para “inferir el constructo subyacente”, permitiendo que el investigador solo pueda asumir causalidad (Wilson, p. 12).

Para asumir una relación causal entre las respuestas observables y el rasgo latente, el desarrollo del instrumento requiere de un proceso riguroso de validación (Cronbach y Meehl, 1955; Messick, 1984). En este caso, un instrumento válido es aquel donde existe amplia evidencia que sugiera que los ítems están midiendo el constructo de interés para la población objetivo. La acumulación de esta evidencia es, en ocasiones, denominada proceso de validación (Shadish, Cook, y Campbell, 2002). Aunque el proceso de validación supone múltiples etapas, un aspecto importante es verificar la estructura correlacional del instrumento en la población de evaluados. Asumiendo que la estructura factorial se mantiene en esta población, el desarrollador del instrumento luego continúa con los siguientes pasos de la validación, como por ejemplo, estudios multirasgo-multimétodo.

Cuando un instrumento está diseñado para ser utilizado en múltiples poblaciones, como es el caso de los ILSAs, se requieren algunos pasos de validación adicionales para verificar que el instrumento opera de igual manera en todas las poblaciones. Brown (2015) menciona cuatro tipos de invarianza grupal para este propósito: a) formas iguales, b) pesos factoriales iguales, c) interceptos/umbrales iguales y d) varianzas residuales iguales (también conocidos como invarianza configural, métrica, escalar y estricta, respectivamente). El modelo de formas iguales es el tipo de invarianza más laxo y supone que la estructura de las relaciones ítem-constructo es idéntica en todos los grupos (ver representación gráfica en cuadrante 1 de la Figura 1). La estructura de pesos factoriales iguales se contrasta con la estructura previa (configural) y requiere que la varianza verdadera de cada ítem sea idéntica en todos los grupos (ver cuadrante

2 de la Figura 1). Luego, el modelo de interceptos iguales, para ítems continuos, y umbrales iguales, en el caso de ítems discretos, supone que los ítems tienen la misma localización en el espacio latente (ver cuadrante 3 en la Figura 1). Finalmente, el modelo de varianzas residuales iguales, cuando se contrasta con el modelo de interceptos/umbrales iguales, indica que todos los ítems tienen la misma varianza en cada grupo, ya que el peso factorial y la varianza residual equivalen a la varianza total. Asegurar la invarianza del modelo de medición implica que el mismo constructo está siendo medido de igual manera en los diferentes grupos. Contar con evidencia de invarianza del modelo de medición no verifica automáticamente la relación causal entre el constructo y las respuestas; sin embargo, la incapacidad de demostrar equivalencia entre las poblaciones sugiere que el supuesto de causalidad no se mantiene. Es importante también mencionar que el nivel de invarianza requerido depende de los objetivos del análisis. Niveles diferentes de invarianza permiten realizar distintos tipos de comparaciones. La Figura 1 muestra un resumen de los tipos de comparación que se pueden realizar en distintos niveles de invarianza.

FIGURA 1. Diferentes niveles de invarianza y tipos de comparaciones permitidas en cada nivel



Nota: las líneas punteadas representan la parte del modelo que se pone a prueba en cada nivel de invarianza.

Método

Datos

Los datos de este estudio provienen de los último ciclos de tres ILSAs: TERCE, implementado por el Laboratorio Latinoamericano para la Evaluación de la Calidad de la Educación de la UNESCO (LLECE); TIMSS, conducido por la Asociación Internacional para la Evaluación del Logro Educativo (IEA) y PISA, realizado por la Organización para la Cooperación y Desarrollo Económicos (OCDE). Estos tres estudios, TERCE, TIMSS 2015 y PISA 2015 son los estudios comparativos internacionales más recientes que evalúan el logro de los estudiantes y recopilan información de un conjunto de actores educativos. Específicamente, PISA mide el logro de los estudiantes de 15 años en matemática, ciencias y lectura, con foco especial en la habilidad del estudiante para aplicar el conocimiento en contextos prácticos y situaciones de la 'vida cotidiana' (OECD, 2014, p. 24). En contraste al foco práctico de PISA, TIMSS y TERCE son instrumentos basados en el currículo y se enfocan en lo que los estudiantes han podido aprender en la escuela. TIMSS mide el logro en matemática de estudiantes de 4to y 8vo grado (Mullis, I.V.S., Martin, M.O., Foy, P., y Hooper, M., 2016), mientras que TERCE mide lectura, matemática y ciencias en 3er y 6to grado (Treviño, et al., 2015). En el presente estudio, utilizamos datos de los 72 sistemas educacionales que participaron en PISA 2015, los 44 sistemas educacionales que participaron en TIMSS 8vo grado, y los 16 sistemas educativos que participaron en TERCE 6to grado.

De cada estudio, seleccionamos intencionalmente una escala que cada organización implementadora construyó y liberó en sus bases de datos como un indicador de antecedentes familiares. Estas escalas provienen de los cuestionarios de estudiantes que se administran a cada participante después de haber completado la parte cognitiva de la evaluación. En PISA, el nivel socioeconómico se estima a través del Índice de Estatus Económico, Social y Cultural (ESCS, por sus siglas en inglés), obtenido a través de un conjunto de variables relacionadas con los antecedentes familiares del estudiante, por ejemplo, educación de los padres, ocupación de los padres, bienes que posee la familia en el hogar, posesiones culturales, y el número de libros disponibles en el hogar (OECD, 2016c, p.205). En TIMSS, utilizamos la Escala de Recursos Educativos en el Hogar (HERS, por sus siglas en inglés), que fue creada a partir de las

respuestas de los estudiantes acerca de la disponibilidad de tres recursos: número de libros en el hogar, nivel educacional más alto alcanzado por los padres y un número de herramientas educativas en el hogar (Martin, Mullis, Hooper, Yin, Foy y Palazzo, 2016). En TERCE, utilizamos la Escala de Nivel Socioeconómico y Cultural de la Familia (ISECF), que proviene de los siguientes ítems: educación de los padres, ocupación de los padres, ingreso familiar y disponibilidad de diferentes posesiones y servicios en el hogar (UNESCO-OREALC, 2016). Debido a que los constructos teóricos de los tres estudios no son equivalentes, el objetivo de nuestro análisis fue examinar el grado en que las organizaciones implementadoras son capaces de crear una escala que sea comparable entre los países que participan en sus estudios, y no comparar la misma escala entre distintos estudios. La Tabla 1 muestra el conjunto de indicadores utilizados en cada estudio para medir los antecedentes socioeconómicos.

TABLA 1. Indicadores utilizados en cada estudio para construir una medida de antecedentes socioeconómicos.

Escala/Estudio	Ítem	Descripción
PISA: Índice de Estatus Económico, Social y Cultural (ESCS) ²	<ol style="list-style-type: none"> 1. Estatus ocupacional más alto de los padres (HISEI). 2. Nivel educacional más alto de los padres (PARED). 3. Posesiones en el hogar (HOMEPOS). 	<ol style="list-style-type: none"> 1. Los datos ocupacionales del padre y la madre se obtuvieron a partir de preguntas abiertas. Las respuestas se codificaron en códigos ISCO de cuatro dígitos y luego fueron agrupadas en un índice socioeconómico internacional de estatus ocupacional. 2. Nivel más alto de educación de cualquiera de los padres, recodificado en las siguientes categorías: (0) ninguno, (1) educación primaria, (2) secundaria baja, (3) secundaria alta vocacional / pre-vocacional, (4) secundaria alta general y/o post-secundaria no terciaria, (5) vocacional terciaria y (6) terciaria teóricamente orientada o postgrado. El índice corresponde al nivel ISCED más alto de cualquiera de los padres. 3. Los estudiantes reportaron la disponibilidad de 16 bienes en el hogar, incluyendo tres bienes específicos de cada país y el número de libros en la casa. Luego, se calculó un índice que resume todos bienes y posesiones del hogar utilizando un modelo TRI mediante WLE (logits) para las dimensiones latentes, las que fueron transformadas a escalas de media 0 y desviación estándar 1 (con pesos muestrales iguales).

⁽²⁾ Estas variables se derivan consecuentemente de un conjunto de ítems individuales. Ver (OCDE, 2016c) para más detalles sobre el procedimiento llevado a cabo para construir esta escala.

<p>TIMSS: Recursos Educativos en el Hogar (HERS)</p>	<p>1. Número de libros en el hogar. (BSBG04). 2. Número de apoyos para el estudio en el hogar. (BSDG06S). 3. Nivel más alto de educación de cualquier de los padres (BSDGED-UP).</p>	<p>1. Categorías de respuesta: (1) 0-10, (2) 11-25, (3) 26-100, (4) 101-200, (5) Más de 200. 2. Categorías de respuesta: (1) Ninguno, (2) Conexión a Internet o habitación propia, (3) Ambas. 3. Nivel educativo más alto de cualquiera de los padres, recodificado en las siguientes categorías: (1) primaria completa o algún grado de secundaria baja o no fue a la escuela, (2) secundaria baja completa, (3) secundaria alta completa, (4) educación post-secundaria completa, (5) universitaria o mayor completa.</p>
<p>TERCE: Escala de Nivel Socio-económico y Cultural de la Familia (ISECF)</p>	<p>1. Nivel educativo más alto de la madre (DQ-FIT09_02). 2. Nivel ocupacional más alto de la madre (DQ-FIT11_02). 3. Ingreso mensual del hogar (DQ-FIT12). 4. Material del piso de la casa (DQ-FIT14). 5. Servicio en el hogar (BIENES1). 6. Posesiones en el hogar (BIENES2). 7. Número de libros en la casa (DQ-FIT21).</p>	<p>1. Categorías de respuesta: (1) ninguno, (2) educación primaria, (3) mayor a educación primaria. 2. Categoría de respuesta: (1) Nunca ha trabajado fuera del hogar, (2) personal de limpieza, mantenimiento, construcción, agricultor, etc., (3) vendedor, operario de máquinas, conduce vehículos motorizados, etc., (4) trabajo administrativo, dueño de un negocio pequeño, (5) profesional, dueño de un negocio grande, a cargo de una división o área de una compañía, etc. 3. Ingreso declarado recodificado en deciles de ingreso del país con las siguientes categorías: (1) decil 1, (2) decil 2, (3) decil 3, (4) decil 4, (5) decil 5, (6) decil 6 a 10. 4. Categorías de respuestas: (1) tierra, (2) cemento o tablas de madera sin pulir., (3) Baldosas, cerámica o similar, (4) Parquet, madera pulida o piso alfombrado. 5. Los estudiantes reportaron la disponibilidad de 5 servicios en el hogar: desagüe o alcantarillado, recolección de basura, teléfono fijo, televisión por cable o satelital y conexión a internet. Luego, se utilizó análisis de componentes principales (ACP) para crear un índice resumen de los servicios en el hogar. 6. Los estudiantes reportaron el número de los siguientes bienes en el hogar: televisor, radio, computador, refrigerador, lavadora de ropa, celular, auto. Luego, se utilizó un análisis de componentes principales (ACP) para crear un índice que resume los ítems bienes en el hogar. 7. Categorías de respuesta: (1) ninguno, (2) 10 o menos, (3) 11-20, (4) 21-30, (5) más de 31.</p>

Fuente: OECD, 2016c; Martin, et al., 2016; UNESCO-OREALC, 2016.

Estrategia analítica

Nuestra estrategia analítica consistió en dos etapas. Utilizamos primero un análisis factorial confirmatorio (AFC) con el objetivo de comprobar la estructura factorial del modelo utilizado para medir algún aspecto de antecedentes socioeconómicos³ en cada estudio (i.e., TIMSS, PISA y TERCE). Se ajustó un modelo AFC para cada país en cada estudio. Luego, utilizamos un análisis factorial confirmatorio multigrupo (AFCMG) para verificar diferentes niveles de invarianza entre los países para cada escala en cada uno de los tres estudios descritos anteriormente. AFCMG (Jöreskog, 1971) es una de las técnicas más utilizadas para evaluar la invarianza de los modelos de medición (Billiet, 2003). AFCMG es una extensión del AFC que se utiliza para evaluar diferencias grupales en medias y covarianzas dentro de un modelo factorial común (Jöreskog, 1971); o como McGrath (2015) señala, para evaluar el ajuste global en múltiples grupos (sistemas educacionales en nuestro caso).

En la primera etapa, para evaluar el ajuste de cada modelo en cada país, utilizamos cuatro indicadores: la prueba de chi-cuadrado, el índice de ajuste comparativo (CFI), el índice de Tucker-Lewis (TLI) y la raíz del error cuadrático medio de aproximación (RMSEA). Adoptamos los puntos de cortes propuestos por Rutkowski y Svetina (2014) para realizar análisis en contextos donde la cantidad de grupos es grande y los tamaños muestrales son grandes y heterogéneos (e.g., muestras de ILSAs): $\leq .10$ para el RMSEA, $\geq .95$ para CFI y TLI. Aunque la prueba de chi-cuadrado no es considerada de utilidad en este contexto (Meade et al., 2008; Rutkowski y Svetina; 2014, Cheung y Rensvold, 2002), también reportamos su valor para analizar si las escalas de comportan de la manera esperada en todas las condiciones, donde los chi-cuadrado generalmente aumenta a medida que se especifican más restricciones en estos modelos.

Es importante comentar que en aquellos casos donde la escala de nivel socioeconómico está compuesta solamente de tres indicadores, como TIMSS y PISA, el modelo de un factor es un modelo saturado (i.e., no tiene grados de libertad). Como consecuencia, la evaluación

³ Aunque la escala de TIMSS no es un índice de nivel socioeconómico, la Escala de Recursos Educativos en el Hogar se utiliza comúnmente en las publicaciones de la IEA como una medida de aproximación a los antecedentes socioeconómicos de los estudiantes. Ver por ejemplo: Martín et al. 2013; Erberber et al., 2015; Trude y Gustafsson, 2016.

del ajuste no se puede realizar porque un modelo de tres indicadores tiene ajuste perfecto. De cualquier modo, de acuerdo con Brown (2015) estos “modelo[s] pueden evaluarse igualmente en términos de su interpretabilidad y del valor de sus parámetros estimados (e.g., magnitud de los pesos factoriales)” (pp. 71).

En la segunda etapa, con el objetivo de verificar la invarianza de las escalas de nivel socioeconómico entre los grupos (i.e., sistemas educativos), se ajustaron modelos AFCMG en todos los grupos simultáneamente dentro de cada estudio. Esto es, se ajustó un modelo AFCMG con todos los países participantes de TERCE, se ajustó un segundo modelo AFCMG para todos los países participantes de TIMSS, y un tercer modelo AFCMG se ajustó para todos los países participantes de PISA. De acuerdo a las prácticas comunes en este campo, ajustamos una serie de modelos anidados que van desde los modelos menos restrictivos a los más restrictivos. De esta manera, comenzamos evaluando cada escala desde el modelo configural, seguido por el modelo métrico y escalar. Aunque es posible evaluar el modelo de invarianza restrictiva o de varianzas residuales iguales (i.e., el cuarto nivel de invarianza) en la jerarquía propuesta por Brown (2015), la invarianza escalar es suficiente para realizar comparaciones significativas entre las medias latentes de los grupos (Marsh et al., 2010; Meredith 1993).

En esta segunda etapa, llevamos a cabo dos conjuntos de análisis. Primero, para examinar el comportamiento de los índices de ajuste de los AFCMG, los modelos se ajustaron para todos los países simultáneamente en cada estudio, donde la prueba de invarianza configural fue seguida por las pruebas de invarianza métrica y escalar. Siguiendo a Rutkowski y Svetina (2014), este primer conjunto de análisis fue denominado *medidas de ajuste global*. Evaluamos cada modelo (e.g., configural, métrico y escala) utilizando los mismos criterios presentados anteriormente. De esta manera, CFI y TLI debieron ser no menores a .95, y RMSEA no mayor a .10.

Luego, para verificar la factibilidad de los modelos de invarianza métrica y escalar, utilizamos ΔCFI , ΔTLI y $\Delta RMSEA$ entre modelos consecutivos más y menos restrictivos (configural vs métrico, y métrico vs escalar). Este segundo conjunto de análisis fue denominado *medidas de ajuste relativo*. Considerando los tamaños muestrales grandes y heterogéneos y el número relativamente grande de grupos (i.e., sistemas educacionales), seguimos la propuesta de Rutkowski y Svetina (2014). Para comprobar

la invarianza métrica, estas diferencias debieron ser $\Delta CFI \leq 0.020$, $\Delta TLI \leq 0.020$ y $\Delta RMSEA \leq 0.030$. Para comprobar la invarianza escalar, las diferencias debieron ser $\Delta CFI \leq 0.010$, $\Delta TLI \leq 0.010$ y $\Delta RMSEA \leq 0.010$.

Resultados

Primero, para cada estudio/escala y para cada país, mostramos los resultados generales de los modelos AFC que describen el grado en que los indicadores empíricos se ajustan a los constructos teóricos propuestos por cada estudio. En segundo lugar, para cada estudio/escala mostramos los resultados del análisis multigrupo y las pruebas de invarianza entre todos los países participantes de cada estudio.

Etapas 1: Análisis por país. Comenzamos nuestro análisis con AFCs para cada país en cada estudio. Como las escalas de TIMSS y PISA tienen solo tres ítems, solo existe un único conjunto de parámetros que se ajustan y reproducen los datos (Harrington, 2009). Por esta razón, en vez de presentar una tabla con los índices de ajuste (que incluiría sólo valores constantes), seguimos el procedimiento propuesto por Miranda y Castillo (2018) y presentamos un gráfico que ilustra los pesos factoriales estandarizados de cada ítem. Esto nos permite evaluar los modelos en relación a la magnitud de los pesos factoriales de cada ítem (Brown, 2015). La Figura 1 presenta los pesos factoriales de la escala de PISA, la Figura 2 contiene los de la escala de TIMSS y la Figura 3 los de la escala del TERCE. Para mantener la consistencia, presentamos el gráfico de pesos factoriales estandarizados del modelo del TERCE, aun cuando este está identificado ($gl > 0$).

En las Figuras 1, 2 y 3, cada punto representa el peso factorial estandarizado de cada ítem en un país determinado, mientras que la línea horizontal que cruza cada punto representa el intervalo de confianza al 95%. Incluimos una línea vertical en el eje que representa un peso factorial de 0.5, ya que este valor se considera el mínimo aceptable para un peso estandarizado en AFC (Hair et al., 2006).

FIGURA 2. Pesos factoriales estandarizadas para cada ítem que compone el SES en PISA, TIMSS y TERCE.

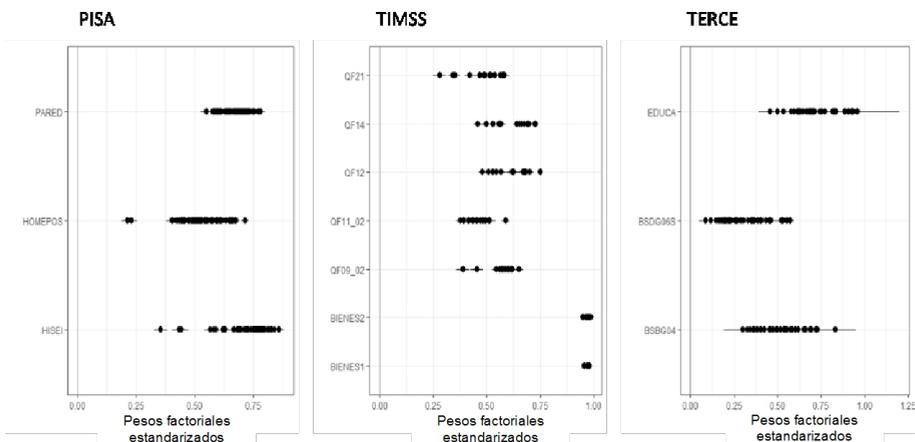


TABLA 2. Número de países con pesos factoriales bajo/sobre 0.5 por indicador/estudio.

Estudio	Indicador	Peso factorial	
		< 0.5	> 0.5
PISA	HISEI	3	65
	PARED	0	68
	HOMPOS	25	43
TIMSS	BSBG04	16	28
	BSDG06S	37	7
	EDUCA	1	43
TERCE	QF09_02	0	16
	QF21	7	9
	QF11_02	12	4
	QF12	1	15
	QF14	1	15
	BIENES1	0	16
	BIENES2	0	16

La Figura 2 presenta los pesos factoriales de cada indicador utilizado para medir antecedentes socioeconómicos en los tres estudios, mientras que la Tabla 3 muestra el número de países que presentan pesos factoriales menores y mayores a 0.5 en cada indicador, para cada estudio considerado. Como se observa, en PISA los indicadores PARED y HISEI muestran pesos factoriales mayores a 0.5 en la mayoría de los países (68 y 65, respectivamente. Ver Tabla 2). El indicador HOMEPOS presenta pesos factoriales bajo 0.5 en 25 países, e incluso valores menores a 0.25 en dos países (ver Tabla 2). En TIMSS, como se observa en la Figura 2, el indicador EDUCA es el único que tiene pesos factoriales mayores a 0.5 en la mayoría de los países (43, ver Tabla 2). Los otros dos indicadores presentan mayor variación entre países. Por ejemplo, el indicador BSBG04 presenta pesos factoriales menores a 0.5 en 16 países, y el indicador BSDG06S presenta pesos factoriales bajo 0.5 en 37 países, con valores bajo 0.25 en casi la mitad de estos (ver Tabla 2). Finalmente, en TERCE, la Figura 2 muestra que ninguno de los indicadores presenta pesos factoriales menores a 0.25. Los ítems QF21 y QF11_02, sin embargo, presenta pesos factoriales menores a 0.5 en algunos países (7 y 12, respectivamente. Ver Tabla 2). Particularmente BIENES1 y BIENES2, presentan pesos factoriales mayores a 0.5 en los 16 países participantes del estudio (ver Tabla 2).

Hasta ahora, hemos ilustrado que las medidas de nivel socioeconómico analizadas y su configuración en los países presentan variaciones importantes entre los estudios. Nuestros resultados sugieren que, entre las escalas de antecedentes socioeconómicos analizadas, la escala de TERCE es la que tiene menor variación en su configuración entre los países y la que tiene mejor ajuste, seguida por las escalas de PISA y TIMSS.

Etapa 2: análisis multigrupo. Las pruebas de invarianza indicaron que se alcanzaron diferentes niveles de invarianza en los tres estudios analizados. En el caso de PISA, a partir de la información entregada por los pesos factoriales, es heurísticamente razonable pensar que la estructura de la escala es similar en todos los países (ver Figura 2). Los tres indicadores tuvieron parámetros relativamente estables entre los países, con pesos factoriales por sobre 0.50 en la mayoría de los indicadores y la mayoría de los países. Sólo el índice HOMEPOS mostró algunos pesos factoriales menores a 0.25 (Qatar y Emiratos Árabes Unidos). Como se observa en la Tabla 3, el modelo métrico mostró índices de

ajuste por sobre los criterios de corte, mientras que el modelo escalar mostró índices de ajuste bajo los criterios establecidos (ver Tabla 3). Sin embargo, las medidas de *ajuste relativo* indicaron que no se alcanzó el nivel de invarianza métrica ni el escalar.

TABLE 3. Medidas de ajuste global de los AFCMG para cada nivel de invarianza

Modelo	PISA				TIMSS				TERCE			
	X ²	CFI	TLI	RMSEA	X ²	CFI	TLI	RMSEA	X ²	CFI	TLI	RMSEA
Configural	0.000	1.000	1.000	0.000	0.000	1.000	1.000	0.000	4124.600	0.979	0.968	0.070
Métrico	5951.480	0.974	0.961	0.077	3195.150	0.863	0.790	0.077	6216.210	0.968	0.965	0.072
Escalar	61932.700	0.728	0.793	0.177	19723.990	0.146	0.672	0.096	16862.100	0.910	0.925	0.107

En TIMSS, la información de ajuste global (e.g., pesos factoriales) muestra que el modelo base de invarianza configural posee una gran dispersión de los pesos factoriales (ver Figura 2). Concretamente, solo en siete países encontramos pesos factoriales sobre 0.05. y cerca de un tercio de los países (e.g., Canadá, Hungría, Irlanda, Italia, Japón y Kuwait) tuvieron pesos factoriales bajo 0.25. Los otros dos indicadores son relativamente estables entre los países. El modelo métrico mostró índices de *ajuste global* por debajo de los criterios de corte y, en consecuencia, no obtuvimos evidencia para afirmar que se alcanzó el nivel de invarianza métrica o escalar (ver Tabla 3). Asimismo, los índices de *ajuste relativo* sugieren que no se alcanzó ni el nivel de invarianza métrica ni el escalar (ver Tabla 4).

Finalmente, el TERCE mostró buenos índices de *ajuste global* en el caso de los modelos configural y métrico, pero estuvieron fuera del rango aceptable en el caso del modelo escalar (ver Tabla 3 y Figura 2 para pesos factoriales del modelo configural). En relación a las medidas de ajuste relativo, las comparaciones entre los modelos configural y métrico entregan evidencia de invarianza métrica (ver Tabla 4).

En resumen, nuestros análisis mostraron índices de ajuste que no entregan evidencia de invarianza métrica o escalar en el caso de las escalas de antecedentes económicos utilizadas en TIMSS y PISA, mientras que la escala del TERCE sí cuenta con evidencia de invarianza métrica y escalar.

TABLA 4. Medidas de ajuste relativo de los AFCMG para cada nivel de invarianza.

Model	PISA			TIMSS			TERCE		
	χ^2 diff.	Δ CFI	Δ RMSEA	χ^2 diff.	Δ CFI	Δ RMSEA	χ^2 diff.	Δ CFI	Δ RMSEA
Métrico	5951.476	0.026	0.077	3126.102	-0.137	0.077	2086.991	-0.011	0.002
Escalar	44692.380	0.246	0.100	17830.425	-0.717	0.019	7523.642	-0.058	0.035

Discusión

A nivel elemental, los cuestionarios de contexto están compuestos por múltiples instrumentos, algunos de ellos diseñados para medir constructos hipotéticos (e.g., estatus socioeconómico). Como estos constructos hipotéticos no se pueden medir directamente, son las respuestas a un determinado conjunto de ítems de los cuestionarios de contexto las que sirven como indicadores indirectos del constructo, el que se operacionaliza a través de un modelo de medición. Como los constructos son teóricos, fenómenos no-observables, el acto de verificar o validar el instrumento es un proceso extremadamente importante que cae dentro del marco de la teoría moderna de la medición (e.g., van der Linden, 2005; Wilson 2005). Las ideas centrales para desarrollar y validar dichos instrumentos son: comenzar con un constructo bien definido, diseñar un conjunto de ítems que se supone eliciten dicho constructo y luego verificar que el modelo de medición propuesto sea consistente con los datos generados por esos ítems. Cuando un modelo propuesto no se ajusta a las respuestas de uno o más ítems, este o estos se deben revisar o reemplazar. En otras palabras, el desarrollo de constructos no debiese ser un ejercicio posterior, en el que se exploran ítems para definir constructos posibles, los que luego se mapean nuevamente en alguna teoría. Al menos, no debiese ser un ejercicio lineal que termina con el mejor intento (aunque insuficiente) para ajustar los datos empíricos a una teoría determinada.

Además, solo luego de que existe evidencia suficiente que sugiera que el instrumento ajusta bien en una población, es que puede comenzar la tarea de evaluar la factibilidad del instrumento para realizar comparaciones entre poblaciones. Un procedimiento común para verificar la invarianza de los modelos de medición entre países es evaluar la equivalencia de la estructura de covarianzas, medias y varianzas residuales. Al examinar el

grado de igualdad del modelo de medición, somos capaces de verificar estadísticamente el supuesto de comparabilidad de las escalas. Si el supuesto se mantiene entonces habría evidencia estadística de que las escalas pueden ser comparadas razonablemente. Pero, como ocurrió en el presente análisis, los constructos no siempre son comparables, lo que sugiere que los atributos medidos no son iguales en los países.

Existe una serie de explicaciones plausibles cuando se descubre que las escalas no son invariantes entre poblaciones o culturas. Lo primero, y lo más importante, es que es absolutamente posible que el constructo teórico no sea pertinente o que haya sido pertinente alguna vez, pero debido a cambios en la sociedad, ocurre que ya no tiene relevancia. Por supuesto, si el constructo no es relevante la escala no debería reportarse y el constructo antiguo debiese reemplazarse por un constructo nuevo. En situaciones donde existe gran apoyo para un constructo universal, existen otras posibles razones por la que una escala puede no ser invariante, incluyendo:

- 1) El constructo se puede medir, pero el marco teórico no es pertinente;
- 2) El constructo se puede medir, el marco teórico es pertinente, pero los indicadores se operacionalizaron incorrectamente.
- 3) El constructo se puede medir, el marco teórico es pertinente, pero no existen indicadores universales.

En relación al primer punto, una estrategia viable y relativamente sencilla de realizar consiste en revisar el marco teórico utilizado para operacionalizar el constructo en cuestión. El segundo punto tiene un poco más de matices. Operacionalizar constructos incorrectamente podría ocurrir por una serie de motivos. Por ejemplo, el marco de referencia que constituye el fundamento para medir el NSE de los niños debiese incluir el ingreso del hogar, una pregunta que es difícil obtener de manera confiable a partir de niños jóvenes. De esta manera, deben recolectarse otros indicadores más indirectos del ingreso del hogar. Aunque obtener indicadores de posesiones en el hogar en estas circunstancias sea lo más razonable, estas variables pueden no reflejar precisamente el ingreso del hogar. Por ejemplo, es posible que, con el auge de la lectura online, el número de libros en el hogar ya no represente ni el capital económico ni el NSE. En el tercer escenario, el constructo existe, pero los indicadores necesarios para medir el constructo difieren entre países o regiones. Nuevamente, el NSE es un ejemplo útil para ilustrar este punto.

La mayoría de la teoría académica define el NSE como un constructo universal. Para fundamentar este constructo, aquellos investigadores interesados en medir el NSE en estudios internacionales podrían aplicar un marco conceptual universal. Pero a pesar de la aceptación de la teoría del NSE, los indicadores que representan el constructo pueden diferir por país. Por ejemplo, un indicador confiable de capital económico familiar en EEUU podría ser si el niño tiene una habitación propia o si la familia ha tenido vacaciones en el extranjero. En cambio, un indicador relativamente deficiente sería si la familia tiene una cortadora de césped. Por el contrario, una cortadora de césped es un signo importante de capital económico en Hong Kong o Singapur, dada la relativa falta de tierras donde hacer crecer el césped. Hasta ahora, la pregunta sobre la universalidad de los indicadores pero no de la teoría se mantiene – existe algún conjunto de indicadores que permita diferenciar internacionalmente y de manera confiable aquellos niños de altos y bajos recursos? Claro es que, dado el comportamiento de las mediciones que se examinaron aquí, así como también otras investigaciones similares (Caro, Sandoval-Hernandez y Lüdtke, 2016; Rutkowski y Rutkowski, 2017), aún queda mucho trabajo por hacer.

Un camino posible es relajar el requerimiento de que los constructos deban ser definidos de manera idéntica en todos los sistemas evaluados. Aunque PISA, por ejemplo, ha permitido la inclusión de ítems de capital económico específicos de cada país, estos no son incluidos en el modelo unidimensional del NSE. En cambio, son tratados como variables observables y específicas del país que luego se transforman en combinaciones lineales de variables para construir una medida de estatus sociocultural. Sin embargo, tales combinaciones lineales no corresponden a variables latentes (Bentler, 1982) y no tienen una estructura hipotetizada. Estas aproximaciones no miden algo, sino que corresponden a meros ejercicios de reducción de datos.

Es posible, en cambio, ajustar modelos de variables latentes que adhieran al supuesto de invarianza parcial, donde se permite incluir ítems únicos y estimar sus parámetros. Investigaciones previas han mostrado que, aunque operacionalizar estos constructos requiere más esfuerzo, esta parece ser una manera promisoría de mejorar la consistencia de los datos de los modelos entre países, manteniendo la comparabilidad.

Rutkowski y Rutkowski (2017) concentraron sus esfuerzos en la región Nórdica, que es relativamente homogénea, y demostraron que se

requiere investigar más para desarrollar medidas únicas de cada país que funcionen adecuadamente. Aunque sin duda esto requerirá un esfuerzo significativo por parte de los países participantes, también permitirá que estos incorporen matices locales y culturales, propios de sus contextos, a escalas posibles de comparar en estudios internacionales.

TERCE entrega otra posible solución. Dado que es una evaluación regional que se centra en países de lenguajes, culturas y economías similares (en comparación a PISA y TIMSS), TERCE debería ser capaz, con mayor dedicación, de diseñar y administrar cuestionarios que se ajusten mejor a poblaciones específicas. En el presente artículo, nuestros resultados indican que el TERCE fue capaz de desarrollar una escala de antecedentes socioeconómicos que es comparable a nivel métrico, lo que supera a su contraparte en TIMSS. A pesar de todo, ningún estudio mostró invarianza escalar, donde las medias latentes pudiesen ser válidamente comparadas entre países. En otras palabras, en TERCE, TIMSS y PISA, existe evidencia estadística para sugerir que el indicador de antecedentes socioeconómicos no es comparable entre distintas culturas. Peor aún, tanto en PISA como en TIMSS, las escalas no cumplen con los estándares básicos de calidad en muchos de los sistemas educacionales participantes. Dado esto, los análisis que utilizan los promedios de las escalas de nivel socioeconómico en cualquiera de estos estudios producen resultados, al menos, cuestionables.

Estos resultados tienen implicancias directas sobre las políticas educativas y la investigación. Por ejemplo, los estudios que estiman la proporción de estudiantes resilientes⁴ en un grupo de países y luego realizan comparaciones entre países son un clásico ejemplo de esta práctica (e.g. OECD, 2011; Erberber, Stephens, Mamedova, Ferguson y Kroeger, 2015). Además, lo mismo puede decirse de cualquier estudio internacional comparativo que use el índice de estatus económico, social y cultural (ESCS) de PISA o la Escala de recursos educativos en el hogar (HERS) de TIMSS como variable de control en un modelo de regresión⁵. Nuestros hallazgos suponen una amenaza importare a la validez de estas escalas y cualquier análisis futuro debería alertar a los lectores de estas amenazas.

⁴) Comúnmente definidos como estudiantes con bajo NSE y algo nivel de logro académico.

⁵) De acuerdo a nuestros resultados, este tipo de comparación serían válidas usando la Escala de Nivel Socioeconómico y Cultural de la Familia (NSCF), ya que esta alcanzó el nivel de invarianza métrica.

Conclusión

Las escalas de los cuestionarios de contexto tienen un rol importante para ayudar a explicar el logro educativo. De hecho, algunas escalas han cobrado vida propia y muchas veces operan fuera de los resultados de logro. Por ejemplo, escalas como *bullying*, compromiso de los estudiantes y compromiso cívico son importantes para las políticas educativas y también interesantes cuando no se relacionan con el logro. Aunque utilizamos el NSE, una escala común a las tres evaluaciones, como ejemplo para este estudio, se puede realizar un análisis similar en cualquier estudio que pretenda utilizar escalas de alguna ILSA para realizar estudios transculturales. Además, como se demostró en este artículo, se quiere más trabajo para mejorar las medidas internacionales. Al menos, las escalas de contexto de los ILSAs requieren un proceso de validación tan riguroso como el de las escalas de logro (OCDE, 2014). Tal proceso contribuiría a prevenir el reporte de escalas que no son comparables entre los países participantes.

Como sugirieron nuestros resultados, al adoptar un foco regional más riguroso en el desarrollo de los cuestionarios, los ILSAs podría mejorar la comparabilidad de determinados constructos, tales como el NSE. Concretamente, dos posibles caminos son: mejorar el desarrollo regional de los cuestionarios y entregar mayor financiamiento para el desarrollo de ILSAs regionales y sus cuestionarios. Podría argumentarse que el Estudio Internacional de Educación Cívica y Formación Ciudadana (ICCS), incluyendo sus módulos regionales, representa el primer camino y el TERCE representa el segundo de estos posibles modelos. En cada caso, sin embargo, un marco conceptual claro, que se exprese en escalas regionales específicas, sigue ausente y necesitaría ser desarrollado en extenso. En el caso de ILSAs de mayor escala como PISA y TIMSS, recomendamos, al menos, diversificar la composición cultural de las partes/instituciones a cargo de los marcos de referencia internacionales. Por ejemplo, el grupo experto que supervisó al comité del marco de referencia y los instrumentos de PISA 2015 incluyó once miembros provenientes principalmente de economías desarrolladas de la OCDE (más de la mitad de EEUU y Alemania). La composición del comité claramente no representaba la composición cultural extremadamente diversa de los participantes de PISA (OCDE, 2016b). Finalmente, en los casos donde los países y grupos de países encuentran que el marco

de referencia está mal especificado, los miembros deberían trabajar con las organizaciones desarrolladoras para realizar ajustes a los marcos de referencia y las escalas. Si eso no es posible, entonces los participantes deberían solicitar a las organizaciones de ILSAs la exclusión del país en cualquier reporte de esa escala. Por supuesto, este compromiso trae consigo un costo alto; sin embargo, el costo de publicar y utilizar malas escalas podría ser aún mayor.

Referencias bibliográficas

- Bentler, P. M. (1982). Confirmatory factor analysis via non-iterative estimation. A fast inexpensive method. *Journal of Marketing Research*, 25A(5), 309-318.
- Billiet, J. (2003). Cross-cultural equivalence with structural equation modeling. En J. Harkness, F. Van de Vijver and P. Mohler (Eds.), *Cross-cultural survey methods* (pp. 247-264). NJ: John Wiley and Sons.
- Brown, T. A. (2015). *Confirmatory Factor Analysis for Applied Research*. New York: Guildford Press.
- Buchmann, C. (2002). Measuring family background in international studies of education: Conceptual issues and methodological challenges. En A. C. Porter and A. Gamoran (Eds.), *Methodological advances in cross-national surveys of educational achievement* (pp. 150-197). Washington, DC: National Academy Press.
- Caro, D. H., Sandoval-Hernández, A. and Lüdtke, O. (2016). Cultural, social, and economic capital constructs in international assessments: an evaluation using exploratory structural equation modeling. *School Effectiveness and School Improvement*, 25(3), 433-450.
- Cheung, G. W., and Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9(2), 233-255. DOI: 10.1207/S15328007SEM0902_5
- Cronbach, L. J., and Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52(4), 281-302.
- Erberber, E., Stephens, M., Mamedova, S., Ferguson, S., and Kroeger, T. (2015). Socioeconomically disadvantaged students who are academically successful: Examining academic resilience cross-

- nationally. *IEA's Policy Brief Series*, No. 5, Amsterdam: IEA. Recuperado de http://www.iea.nl/policy_briefs.html
- Gaviria Soto, J. L., Biencinto López, M. C., and Navarro Asencio, E (2007). Invarianza de la estructura de covarianzas de las medidas de rendimiento académico en estudios longitudinales en la transición de Educación Primaria a Secundaria. *Revista de Educación*, 348, 153-173.
- Glas, C., and Jehangir, K. (2014). Modeling country-specific differential item functioning. En L. Rutkowski, M. von Davier, and D. Rutkowski (Eds.), *Handbook of international large-scale assessment: Background, technical issues, and methods of data analysis*. Boca Raton, FL: Chapman and Hall / CRC Press.
- Goldstein, H. (2004). International comparisons of student attainment: Some issues arising from the PISA study. *Assessment in Education: Principles, Policy and Practice*, 11(3), 319–330. DOI: 10.1080/0969594042000304618
- Harrington, D. (2009). *Confirmatory Factor Analysis*. Oxford: Oxford University Press.
- Jöreskog, K. G. (1971). Simultaneous factor analysis in several populations. *Psychometrika*, 36(4), 408-426.
- Kreiner, S., and Christensen, K. B. (2014). Analyses of model fit and robustness. A new look at the PISA scaling model underlying ranking of countries according to reading literacy. *Psychometrika*, 79(2), 210–231. DOI: 10.1007/s11336-013-9347-z
- Martin, M.O., Mullis, I.V.S., Hooper, M., Yin, L., Foy, P. and Palazzo, L. (2016). Creating and Interpreting the TIMSS 2015 Context Questionnaire Scales. En M. O. Martin, I. V. S. Mullis, and M. Hooper (Eds.) *Methods and Procedures in TIMSS 2015*. Chestnut Hill, MA: TIMSS and PIRLS International Study Center / IEA.
- Marsh, H.W., Ludtke O., Muthen, B., Asparouhov, T, Morin, A.J.S., et al. (2010). A new look at the big five factor structure through exploratory structural equation modeling. *Psychol. Assess.* 22(3), 471–91.
- McGrath, R. E. (2015). Measurement Invariance in Translations of the VIA Inventory of Strengths. *European Journal of Psychological Assessment*. On-line advanced publication. DOI: 10.1027/1015-5759/a000248
- Mazzeo, J., and von Davier, M. (2009). Review of the Programme for International Student Assessment (PISA) test design: Recommendations for fostering stability in assessment results. *Education Working Papers EDU/PISA/GB*. Paris: OECD.

- Messick, S. (1984). The Psychology of Educational Measurement. *Journal of Educational Measurement*, 21(3), 215–237.
- Miranda, D. and Castillo, J. C. (2018). Measurement model and invariance testing of scales measuring egalitarian values in ICCS 2009. En A. Sandoval-Hernandez, M. M. Isac and D. Miranda (Eds.) *Teaching Tolerance in a Globalized World*. Cham: Springer International Publishing
- Mislevy, R. J., Beaton, A. E., Kaplan, B., and Sheehan, K. M. (1992). Estimating population characteristics from sparse matrix samples of item responses. *Journal of Educational Measurement*, 29(2), 133–161.
- Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika* 58(4), 525–43
- Mullis, I.V.S., Martin, M.O., Foy, P., and Hooper, M. (2016). *TIMSS 2015 International Results in Mathematics*. Chestnut Hill, MA: TIMSS and PIRLS International Study Center / IEA.
- OECD. (2011). *Against the Odds: Disadvantaged Students who Succeed in School*. Paris: OECD Publishing.
- OECD. (2012). *PISA 2009 Technical Report*. Paris: OECD Publishing.
- OECD. (2014). *PISA 2012 Technical Report*. Paris: OECD Publishing.
- OECD. (2016a). *PISA 2015 Assessment and Analytical Framework*. Paris: OECD Publishing.
- OECD. (2016b). PISA 2015 Background questionnaires. Annex A (pp. 129–196). En *PISA 2015 Assessment and Analytical Framework*. Paris: OECD Publishing.
- Oliveri, M. E., and Ercikan, K. (2011). Do different approaches to examining construct comparability in multilanguage assessments lead to similar conclusions? *Applied Measurement in Education*, 24(4), 349–366. DOI: 10.1080/08957347.2011.607063
- Oliveri, M. E., and von Davier, M. (2011). Investigation of model fit and score scale comparability in international assessments. *Psychological Test and Assessment Modeling*, 53(3), 315–333.
- Oliveri, M. E., and von Davier, M. (2014). Toward increasing fairness in score scale calibrations employed in international large-scale assessments. *International Journal of Testing*, 14(1), 1–21. DOI: 10.1080/15305058.2013.825265
- Rutkowski, L. and Rutkowski, D. (2010). Getting it “better”: The importance of improving background questionnaires in International

- Large-Scale Assessment. *Journal of Curriculum Studies*, 42(3), 411–430. DOI: 10.1080/00220272.2010.487546
- Rutkowski, L. and Rutkowski, D. (2017). Improving the comparability and local usefulness of international assessments: A look back and a way forward. *Scandinavian Journal of Educational Research*, 0(0), 1–14. DOI: 10.1080/00313831.2016.1261044
- Rutkowski, L., Rutkowski, D. and Zhou, Y. (2016). Parameter estimation methods and the stability of achievement estimates and system rankings: Another look at the PISA model. *International Journal of Testing*, 16(1), 1–20.
- Rutkowski, L., y Svetina, D. (2014). Assessing the hypothesis of measurement invariance in the context of large-scale international surveys. *Educational and Psychological Measurement*, 74(1), 31-57.
- Schulz, W., Ainley, J. and Fraillon, J. (Eds.). (2011). *ICCS 2009 Technical Report*. Amsterdam: International Association for the Evaluation of Educational Achievement.
- Shadish, W. R., Cook, T. D., and Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin Company.
- Treviño, E., Fraser, P., Meyer, A., Morawietz, L., Inostroza, P. and Naranjo, E. (2015). *Informe de Resultados TERCE. Factores Asociados*. Santiago: Oficina Regional de Educación para América Latina y el Caribe (OREAL/UNESCO).
- UNESCO-OREALC. (2016). *Reporte Técnico. Tercer Estudio Regional Comparativo y Explicativo, TERCE*. Santiago: Oficina Regional de Educación para América Latina y el Caribe (OREAL/UNESCO).
- Van Der Linden, W. J. (2005). A Comparison of Item-Selection Methods for Adaptive Tests with Content Constraints. *Journal of Educational Measurement*, 42(3), 283-302.
- Wilson, M. (2005). *Constructing Measures: An Item Response Modeling Approach*. Mahwah, NJ: Lawrence Erlbaum Associates.

Dirección de contacto: Andrés Sandoval-Hernández. University of Bath, Faculty of Humanities & Social Sciences, Department of Education. University of Bath, Claverton Down, Bath, BA1 6TP, United Kingdom. **E-mail:** A.Sandoval@bath.ac.uk

Educación y diplomacia cultural en la primavera de Europa (1948-1954)¹

Education and Cultural Diplomacy in the Spring of Europe (1948-1954)

DOI: 10.4438/1988-592X-RE-2019-383-401

Luís Domínguez Castro
José Ramón Rodríguez-Lago
Universidad de Vigo

Resumen

El papel que jugó la escuela y la educación en los debates durante la denominada “primavera de Europa” no ha sido suficientemente valorado. Instituciones como la Unión Occidental, la Unesco, el Movimiento Europeo o el Consejo de Europa organizaron un buen número de seminarios y encuentros de maestros, profesores, inspectores e intelectuales que reflexionaron sobre cómo formar a la ciudadanía en un marco de paz y cooperación internacional y cómo introducir la idea de Europa desde la escuela. Aplicando los principios teóricos del poder blando y relacional de Nye y utilizando una metodología cualitativa e historiográfica, analizamos la documentación conservada de las reuniones para ver qué recursos emplea cada Estado en la lucha por la hegemonía cultural en el ámbito educativo y para acordar los principios inspiradores de la idea de Europa a transmitir desde la escuela. Las lenguas modernas y la historia van a ser las disciplinas que más atención reciban. Se afirma que es necesaria la presencia de una lengua extranjera en el currículo de primaria. El inglés avanza como lengua escolar y el francés se resiste a perder su hegemonía a través de una diplomacia cultural muy activa. Se rechaza elaborar un manual de historia general de Europa,

¹ Este trabajo fue financiado con cargo a los proyectos HAR2015-64429-C2-1-P del Ministerio de Economía y Competitividad y JEAN MONNET PROJECT, 574846-EPP-1-2016-1-ES-EPPJMO-PROJECT de la Education, Audiovisual and Culture Executive Agency. European Commission.

pero se insiste en la necesidad de mitigar los excesos de la mitología nacional en sus páginas. Se considera por último que la historia reciente, debe tener cabida en el currículo escolar, para sofocar los riesgos de su posible manipulación entre las jóvenes generaciones de europeos.

Palabras clave: Escuela, poder blando, poder relacional, educación, diplomacia cultural.

Abstract

The role played by schools and education in the so-called European spring has not been sufficiently assessed. Institutions such as the Western European Union, the Unesco, the European Movement or the Council of Europe have organized a great number of seminars and meetings attended by primary and secondary teachers, education inspectors and intellectuals who discussed how to teach the values of peace and international cooperation and how to introduce the idea of Europe at schools. By applying the theoretical principles of Nye's soft & relational power and using a qualitative and historiographic methodology, we aim at analysing those meetings documents in order to identify the resources are used by each state in their fight for cultural hegemony in education, and to lay the foundations for Europe to be taught at schools. Thus, Modern Languages and History are the most studied disciplines. It is also stated that the teaching of a foreign language is necessary in primary school curricula; English leads the way as schooling language and French fights for keeping its hegemony resorting to a highly active cultural diplomacy. A general European history book won't be warmly received, but to lessen the excesses of national myths in text books is said to be indispensable. Finally, it is considered that recent history must own a place in the curricula so as to avoid the risks of its possible manipulation by European younger generations.

Key words: School, soft power, relational power, education, cultural diplomacy.

Introducción: fundamentación teórica

Los años de posguerra que siguieron a la Segunda Guerra Mundial vinieron marcados por la Guerra Fría, un período de enfrentamiento entre los dos bloques liderados, respectivamente, por EE. UU. y la URSS. En Europa Occidental, esta confrontación se libró intensamente en el ámbito de la

cultura, sin escatimar recursos tanto financieros como organizativos. Este trabajo se enmarca en la llamada *Guerra Fría Cultural* (Saunders, 1999), término que hace referencia al enfrentamiento, en todos los ámbitos, contra la influencia cultural comunista en parte de la *intelligentsia* occidental (Judt, 2007). La educación escolar no ha merecido espacio en la obra ancilar de Saunders ni en monografías posteriores (Scott-Smith y Krabbendam, 2004) centradas en la llamada *alta cultura* –artes plásticas, música clásica, literatura– y en la *cultura de masas* –artes visuales y medios de comunicación–. En consecuencia, pretendemos contribuir a llenar ese vacío demostrando la importancia de los debates sobre educación por parte de cuatro organizaciones muy destacadas en este período, como son la Unión Occidental (UO), la Unesco, el Consejo de Europa y el Movimiento Europeo.

El final de la década de 1940 y los comienzos de la siguiente coinciden con la *primavera de Europa* (Bossuat, 1994 y 2005). Durante esos años, la cultura gozó de cierto protagonismo, si bien otros aspectos tales como el éxito de las organizaciones de integración económica o los fracasos de las de cooperación política y militar han eclipsado los debates culturales de entonces y, entre ellos, los celebrados en torno a la escuela y a su potencial como elemento de socialización de una ciudadanía comprometida con la paz y la concordia internacionales (Gavari Starkie, 2015). Bossuat acaba situando la *Primavera de Europa* entre 1945 y 1949 porque en ese año se frustran las uniones aduaneras impulsadas desde la OECE, se pierde la oportunidad de un Consejo de Europa en clave supranacional y la OTAN ejemplifica la dependencia de los EE. UU. Nosotros optamos por una periodicidad diferente y situamos esa *primavera europea* entre 1948 (cuando los diferentes movimientos de unidad europea se articulan en el Movimiento Europeo y nacen organizaciones como la UO o la OECE) y 1954 (cuando el fracaso de la Comunidad Europea de Defensa despierta de su sueño a quienes imaginaban cercanos los Estados Unidos de Europa).

Un referente teórico básico en este trabajo será el de *poder relacional* (Nye, 2004 y 2011), entendido como la capacidad de convertir en poder real el potencial que supone la posesión de recursos. Los actores transnacionales –como el Movimiento Europeo, la UO, la Unesco y el Consejo de Europa– ejercen un papel central en este nuevo paradigma de las relaciones internacionales, sin menoscabo del desempeñado por los Estados, como organizaciones capaces de construir y mantener redes de

confianza que permitan trabajar en colaboración por metas comunes. El poder relacional es siempre un poder *con* los otros y no *sobre* los otros; así pues, podemos decir que su influencia descansa en el empleo hábil del *poder blando* y del *poder inteligente* para establecer las preferencias de los otros en línea con las del Estado o actor que hace uso de ellos (Nye, 1990 y 2004). Como quiera que tres de las organizaciones con las que trabajaremos son intergubernamentales, conviene recordar que los Estados disponen de tres recursos intangibles con los que ejercer el poder blando: su cultura, sus valores políticos y su política exterior. Ahora bien, tal y como afirma Nye (2004), la diplomacia pública solo es efectiva cuando la cultura del actor que interviene es atractiva para los otros actores, cuando sus valores políticos son respetados por los demás porque se cumplen en el país de origen y cuando su política exterior goza de legitimidad y autoridad moral. La diplomacia pública puede entenderse como el conjunto de acciones mediante las cuales los gobiernos se dirigen a la ciudadanía de otro país con la intención de controlar las agendas públicas de este, modificando sus preferencias al influir en sus expectativas sobre lo que es legítimo y posible realizar. Una de las vertientes de esta nueva diplomacia, por oposición a la tradicional, es la *diplomacia cultural*, que pretende difundir las expresiones culturales de una sociedad determinada, con un destacado papel para la educación.

Los intelectuales promotores de la unidad europea creían que el sistema educativo tenía que permitir la identificación de élites capaces de difundir y concretar la idea europea (Mouvement Européen, 1949). Buena parte de esos intelectuales pertenecía, tanto por edad como por cultura política, a las generaciones defensoras de una educación para las élites, llamadas a transformar desde arriba la sociedad sin arrostrar los peligros inherentes a las transformaciones desde la base. No obstante, la emergencia electoral del laborismo británico y de la democracia cristiana continental, así como la consolidación de la socialdemocracia escandinava en una época en la que el interés de las élites cultas por el comunismo se halla en su momento álgido, hacen preciso que se preste más atención a la educación para las masas, especialmente durante las etapas iniciales de la enseñanza y el acceso a la secundaria. Estamos en los años de introducción de la escuela comprensiva en Suecia, tras algunas experiencias previas en otros países nórdicos y Reino Unido (García Pedraza, 2015).

Por otra parte, la escuela fue uno de los más importantes agentes de nacionalización –junto con el ejército– en la Europa de la segunda mitad del siglo XIX y la primera del XX. No debe sorprendernos, pues, que también los europeístas reparen en ella como factor esencial en el triunfo del nuevo evangelio de unidad que se impartía desde múltiples púlpitos a un lado y a otro del Atlántico. Sin embargo, a la hora de construir una identidad europea y favorecer un clima de concordia y entendimiento a nivel internacional, no todas las materias del currículo escolar tienen el mismo peso. Así, las lenguas modernas y la historia son las que reciben una mayor atención; la diplomacia cultural de los países occidentales no sólo tiene que enfrentar el predicamento del universo comunista, sino también los efectos de la descolonización y la entrada en escena del denominado *Tercer Mundo*. Francia acumulaba una aquilatada experiencia en el uso de la diplomacia cultural (Haigh, 1974). Vamos a intentar comprobar hasta qué punto lo anterior es aplicable también al ámbito educativo, qué otros países entraron en liza y qué recursos de poder blando, inteligente y relacional puso en juego cada uno. Para ello, con el fin de analizar los resultados de los diferentes seminarios, simposios y conferencias que el Movimiento Europeo, la UO, la Unesco y el Consejo de Europa organizaron en torno a las dos disciplinas mencionadas, recurriremos a una metodología de carácter cualitativo (Flick, 2004) que combinaremos con los instrumentos teóricos aportados por Nye, adaptándolos a un escenario anterior a sus formulaciones, como el de la primavera de Europa².

La lucha por la hegemonía cultural en el terreno educativo

Francia, al igual que Alemania, venía desplegando una activa diplomacia cultural a través de la enseñanza, el arte y la arqueología desde finales del siglo XIX. EE. UU. se sumó a esta iniciativa en los años treinta del siglo XX –especialmente en América Latina– con su intento de desplazar al francés como lengua en las escuelas, las propias universidades y los *colleges* americanos sustituyéndolo por el español (Delgado Gómez-Escalonilla, 2015). Durante la posguerra asistimos a la disputa por la

⁽²⁾ Las fuentes documentales proceden de los Archives Historiques de l'Union Européenne (AHUE) en el Instituto Europeo de Florencia y los archivos online del Consejo de Europa y de la Unesco.

hegemonía cultural de Europa entre una Francia heredera de un lugar de privilegio ahora amenazado y una Gran Bretaña con sus velas hinchadas ante la expansión universal de su lengua, en buena parte debido al liderazgo de los EE. UU. en el bloque occidental. En efecto, Jean-Paul Sartre (1949, p.1) sostenía que “*les idées culturelles, tout à fait indépendamment de leur valeur interne, ont un potentiel de diffusion qui dépend de l'importance économique ou militaire du pays considéré... Si nous voulons que la culture française reste, il faut qu'elle soit intégrée aux cadres d'une grande culture européenne*”. En las cinco páginas del documento se menciona a Francia, Alemania, Italia, Holanda, Suecia, Luxemburgo, Grecia, Portugal..., pero ni una sola vez a Gran Bretaña como miembro de esa comunidad cultural europea. Esta sentencia de Sartre encaja muy bien con la noción de *poder inteligente* de Nye, al combinar elementos de *poder duro* –milicia, economía– con elementos de *poder blando* –cultura– y adelantar la necesidad de actuar mediante un poder relacional que permita contar con el apoyo de otros países frente al empuje británico, en beneficio de los intereses franceses.

La educación se convierte en uno de los campos en que se libra la batalla por la hegemonía cultural. Así, el Comité Cultural de la UO, organizó tres ciclos de cursos, cada uno de los cuales estaba destinado a un colectivo distinto: inspectores de educación, maestros o funcionarios gubernamentales. La secuencia siempre fue la misma. Los primeros cursos tuvieron lugar en Gran Bretaña en 1949; los segundos, al año siguiente en Francia. El más importante para nosotros, el destinado a los maestros, comenzó en 1949 en la ciudad inglesa de Ashridge, continuó en 1950 en la francesa Sèvres y concluyó en 1951 en la holandesa Oosterbeek. El resultado de los encuentros anteriores fue la publicación del folleto *The Civilisation of Western Europe and the School*, que, escrito por el pedagogo francés Pierre Joulia y editado en inglés, francés y holandés, contenía toda una serie de sugerencias para que los maestros abordaran esta temática en sus clases (Haigh, 1974).

Pocos días después de la creación de la UO, Ifor Evans –a la sazón rector del Queen Mary College y anteriormente director general de educación del British Council (1940-1944)– le escribe una carta a Kenneth Lindsay –diputado independiente y Ministro de Educación (1937-1940)–, en la que alude al papel que debe representar la cultura, y muy especialmente la educación, en el proceso de unidad europea. Así, Evans (1948) cree

que la escuela primaria es el momento ideal para iniciar el conocimiento mutuo:

“I should like to see prepared in the first place a book for children who are beginning to read which would have the names of the different countries of the Western Union with some of their customs and stories, so that at the age of ten or eleven boys and girls would be as familiar with the names of the members of the Western Union as they are with the words English, Irish, Scots and Welsh. The book I have in mind would be a standard book, but it would be translated into each of the languages of the Western Union and attempts would be made to get the Ministries of Education, Religious Education bodies, etc., to use it.”

En el fondo, se trataba de exportar el modelo de la Union Jack al continente: no en vano el United Europe Movement, impulsado por Churchill, tenía una rama británica y otra francesa, en línea con la Unión franco-británica que el propio líder *tory* había planteado en junio de 1940. Los tres gobiernos del Benelux habían pasado la guerra en el exilio londinense, donde se había ejercido una intensa diplomacia cultural y donde el poder relacional de Gran Bretaña había ocupado un lugar central a través de la Conferencia de Ministros Aliados de Educación (CAME, por sus siglas en inglés).

Para la etapa secundaria, Evans (1948) apunta otra estrategia:

“There would be the problem of affecting children at a somewhat more advanced school age and here I would suggest the preparation of two volumes. One would be a very simple Western Union handbook which would give information on the nature of the constitutions, size, government, etc., of each of the countries in the Western Union. It would give something about their resources and the ways in which they could help each other. It would tell of their school life and their Universities and have some reflections upon their cultural achievements in the past. Once again the book would be a standard book, but it would be translated into each of the languages of the Western Union. Its contents would be agreed contents, though I think once again as we are free nations that a certain amount of compromise and variation should be allowed. The other book I would have in mind would be a little more difficult to prepare. It would be a selection from some of the literature

of each of the countries of the Western Union. I think it might have to be confined in the first instance to only some ten extracts, verse or prose, and they would have to be translated as would be the other projects I have outlined...

Conviene recordar que, en los cinco países de la UO, la educación secundaria era una vía selectiva para forjar élites. Por este motivo, cobra sentido la intención de crear una comunidad de pertenencia e inculcar la cooperación ante un futuro incierto a causa de las amenazas externas. No obstante, la historia, la más destacada de las disciplinas escolares a la hora de construir identidades, no aparece mencionada en las reflexiones de Evans salvo por una referencia a los hitos culturales alcanzados por cada país en el pasado. Sin embargo, sí se destaca otra fuente básica en la conformación de identidades y estereotipos, como es el caso de la literatura (López Facal, 2010).

En paralelo a este temprano interés británico por la educación como base esencial sobre la que elaborar una narrativa europea, la sección francesa del Movimiento Europeo presenta un informe sobre el papel de la educación, a través de Jean Bayet, titular de la cátedra de lengua y literatura latinas en la Sorbona y antiguo Director General de Enseñanza en el gobierno de De Gaulle de 1944, en la Conferencia Europea de la Cultura de Lausana (1949). En su informe, que constituirá la base de las conclusiones de la Conferencia para crear una enseñanza europea a todos los niveles, Bayet acude a claros elementos de poder blando para justificar el derecho francés a establecer las preferencias que serán asumidas por los representantes de los otros países: *“le rayonnement international permanent de plusieurs des Universités françaises, la forte organisation de l’Enseignement d’État en France, son caractère non-confessionnel, les études préliminaires faites à partir de 1945, en vue d’une réforme générale et profondément démocratique de l’enseignement.”* (Bayet, 1949, p. 1).

Abundando en el prestigio del sistema público de enseñanza francés, el informe Bayet defiende, y la Conferencia asume, que en los últimos años de primaria debe introducirse la enseñanza de una lengua viva extranjera, con especial atención a la enseñanza oral y la adquisición de vocabulario. Esta recomendación deriva de la apuesta francesa por la metodología directa en la didáctica de las lenguas vivas, presente ya desde 1902 e inspirada en la Alemania de Bismarck, privilegiando los

métodos orales y activos sobre las tradicionales traducciones escritas heredadas de la enseñanza del latín y el griego (Puren, 2012).

El prestigio de la metodología activa francesa en la enseñanza de las lenguas no sólo dejó su impronta en la Conferencia Europea de la Cultura de Lausana, sino que también hizo lo propio en el seminario que la Unesco dedicó a las lenguas modernas como vía hacia el entendimiento internacional, celebrado en Nuwara Eliya (Sri Lanka) entre el 3 y el 28 de agosto de 1953. En efecto, el documento elaborado por la sección francesa, que apostaba por la metodología activa en la didáctica de las lenguas, fue la base de la preparación del seminario (Morales Gil, 2007). Sin embargo, en este foro los angloamericanos contraatacaron con la defensa de su método audio-oral o militar, recurriendo al prestigioso manual de inglés básico de Ogden y haciendo hincapié en la ausencia de obras similares en las demás lenguas. Los expertos israelíes y australianos presentes corroboraron el éxito de esta metodología aplicada a la rápida inserción de los migrantes (UNESCO, 1954). El resultado fueron unas resoluciones de consenso que combinaban aspectos de ambas metodologías.

Una vez que el Consejo de Europa se inclina a defender la necesidad de introducir el estudio de una lengua extranjera en la escuela, incluido el nivel primario, la disputa se traslada a las lenguas que deberían ser seleccionadas. En 1952, Francia presenta una propuesta para instituir una comunidad lingüística europea basada en el bilingüismo francés-inglés, argumentando que se trata de dos lenguas de amplia expansión y solicitando la inclusión de una de ellas en la etapa de educación primaria de aquellos países que no las tuviesen como propias. De esta forma, el alumnado podría incorporar posteriormente en secundaria otra lengua más –el alemán o el italiano– a su formación. El representante alemán se opone y sostiene la necesidad de un sistema trilingüe en el que se incluya la lengua germana, hablada por 100 millones de europeos de diversas nacionalidades de la Europa central y oriental, amén de ser la lengua de Lutero, Goethe y Heine (Consejo de Europa, 1952b). La disputa entre el francés y el inglés por la hegemonía tuvo episodios poco diplomáticos en el seno del Consejo de Ministros del Consejo de Europa, como cuando el comité de expertos culturales enmendó el acta de uno de sus subcomités, que decía “*Au tours de ces dernières années, un grand nombre d’adultes de certains pays ont appris l’anglais qu’ils n’avaient pas appris à l’école*” (Consejo de Europa, 1952a, p. 13), traducéndola al inglés de manera mucho más imprecisa: “*During the last few years,*

great numbers of adults in some countries have learned languages which they never learnt at school" (Consejo de Europa, 1952a, p. 4). Todo un ejemplo de batalla librada en el terreno del poder blando. Finalmente, el Consejo optó por mantenerse al margen y dejar la decisión en manos de cada Estado miembro, argumentando que los sistemas democráticos no permiten ese tipo de imposiciones y que puede resultar más útil conocer la lengua del vecino que una de gran difusión internacional (Consejo de Europa, 1952a).

Por su parte, la Unesco también optó por no hacer recomendaciones de lenguas concretas, máxime teniendo en cuenta los vestigios del imperialismo cultural, puesto de manifiesto por los expertos asiáticos, que dificultan la empatía del alumnado hacia determinadas lenguas occidentales, amén de que para un oriental su estudio suponía una ruptura con su continuo cultural que entroncaba con el sánscrito, pali o persa y no con el latín o el griego (UNESCO, 1954).

Si en el ámbito de las lenguas modernas Francia y Gran Bretaña/EE. UU se disputan la hegemonía, en el campo de la historia hay más países involucrados. En efecto, los escandinavos, en un claro ejemplo de poder relacional en torno al Consejo Nórdico, van a reivindicar su anticipación en la creación de comisiones bilaterales de revisión de los prejuicios nacionales en los manuales escolares. Ya en la década de 1920 se habían creado comités de historiadores daneses, noruegos y suecos para llevar a cabo esta tarea. El noruego, por ejemplo, pudo constatar como en los libros de texto de la escuela primaria se narraba el episodio de la utilización de la vieja catedral de Trondheim como establo para los caballos de los invasores suecos (1563-1570). Este hecho cobra mayor importancia si se tiene en cuenta que en esa catedral reposan los restos del santo patrón de Noruega, Olaf, y que se reconstruyó como santuario nacional noruego en el siglo XIX. Aun suponiendo que fuese cierto el relato, para conseguir un mayor impacto en la imaginación de los niños se ocultaba que, por aquellos años, la catedral estaba abandonada y en ruinas (Vigander, 1950). Esta experiencia permitió que el seminario Unesco, celebrado entre el 12 de julio y el 23 de agosto de 1950 en Bruselas con el propósito de mejorar los manuales escolares –con especial atención a los de historia–, apoyase la extensión del modelo nórdico de revisión mutua de los manuales con comisiones bilaterales que integrasen a autores, a editores, a maestros y a académicos (UNESCO, 1950). Asimismo, hizo posible que el noruego Haakon Vigander fuese

cooptado a la presidencia del simposio sobre la idea de Europa en la enseñanza de la historia, celebrado entre el 4 y el 12 de agosto de 1953 en Calw (en plena Selva Negra alemana) bajo el patrocinio del Consejo de Europa.

Alemania –al igual que en el terreno lingüístico– tampoco se quedó al margen. Más bien al contrario, su diplomacia cultural se involucró especialmente en el campo de la enseñanza de la historia. La elección de Calw como sede para la celebración del simposio del Consejo de Europa tiene mucho que ver con el proceso de rehabilitación del nuevo estado federal –todavía no soberano–, muy preocupado por el tema de la revisión de los manuales escolares. No en balde el país germánico está presente en ocho de las trece comisiones bilaterales de revisión creadas y, además, se vuelca, a través del Internationales Schulbuchinstitut Braunschweig –hoy sede del Georg-Eckert-Institut für internationale Schulbuchforschung–, en la organización de encuentros sobre el tema celebrados anualmente entre 1951 y 1963. Es más, el simposio de Calw contó con el apoyo explícito del Secretario de Estado de Asuntos Exteriores, Walter Hallstein, que posteriormente se convirtió en el primer presidente de la Comisión Europea. Significativamente, las tres ponencias sobre elementos de una concepción europea de la historia son encomendadas a un alemán –Edad Media–, un francés –Edad Moderna– y un británico –Edad Contemporánea– (Consejo de Europa, 1953b).

En el seno de la Unesco, tal y como hemos comprobado en el ámbito de las lenguas, la enseñanza de la historia también suscitó un importante debate en torno a la hegemonía cultural occidental. Los seminarios de Bruselas de 1950 y de Sèvres de 1951 constataron la necesidad de deseuropeizar los manuales de historia universal. Los argumentos expuestos a favor de la incorporación de otras civilizaciones son un ejemplo *ante* de poder inteligente (UNESCO, 1951b, p. 8): “*Non-European peoples must be given a place which is proportionate to their world importance and their contributions to civilization. Nor should this importance be measured in terms of territorial size[...] military strength or wealth alone; contributions to humanity’s common heritage must be also taken into account.*”. No obstante, este criterio hace que al final se apueste expresamente por una mayor presencia de las civilizaciones india y china, sin mencionar otras en las conclusiones, lo que indica el escaso poder relacional de las naciones latinoamericanas en la Unesco

a pesar de que en aquel momento suponían un tercio de los países miembros.

La escuela como factor de europeización

En el informe para la Conferencia Europea de la Cultura, Jean Bayet (1949) sostiene que generar y difundir un “*esprit européen*” no es tarea de la escuela porque quienes lo aseguraron en la Edad Media no fueron tanto las universidades como las relaciones entabladas por maestros y alumnos venidos de todos los países a ellas, insistiendo en la necesidad de fomentar los intercambios escolares como elemento de poder relacional. Tras los desastres morales y vitales de la guerra y de la posguerra, los mentores de la Conferencia creían que ninguna reorganización política o social podría contar con una base sólida si no se asentaba en una reestructuración radical de la educación y de la enseñanza que debería dominar la formación de la juventud de todos los países de Europa, algo que se había visto dificultado en los últimos siglos por la pérdida de una lengua culta común, por la autonomía y riqueza de las culturas nacionales y por el creciente peso de las concepciones utilitaristas de la cultura y la educación frente a aquellas basadas en el perfeccionamiento espiritual sin objetivo práctico inmediato.

La fuerte impronta federalista de los principales organizadores de la Conferencia –Rougemont, Silva, Madariaga– se traduce en un decidido posicionamiento a favor de la mitigación de los nacionalismos, tan presentes en la formación escolar: “*les éducateurs de tous pays soient préparés à assumer la tâche primordiale de former les nouvelles générations dans un esprit non seulement national mais aussi européen et largement humaniste*” (Mouvement Européen, 1949, p. 15). No obstante, el informe Bayet (1949, p. 2) alertaba sobre el inconveniente de que un ideal de incierta apariencia, como el europeo, dominase de manera prematura y precipitada las sólidas contribuciones de las culturas nacionales:

“Ce n’est qu’à force de siècles, que les provinces françaises se sont groupées en Nation sans perdre trop de leur potentiel propre; une Europe fondée très vite sur une “Koiné” indistincte et sans nuances aurait trop perdu de sa substance et de ses énergies pour être autre

chose qu'une fragile illusion de surface déguisant le néant d'une ruine."

Rougemont (1949), consciente de ello, insiste en que no se pretende, en ningún caso, sustituir veinte nacionalismos locales por un nacionalismo europeo. Europa ha estado, desde sus orígenes, abierta al mundo. En efecto, ha concebido siempre su civilización como un conjunto de valores universales. Se trataba, pues, de que los futuros formadores actuasen como una mancha de aceite sobre una atmósfera cada vez más preparada para recibir su mensaje humanista, como antes lo había estado para interiorizar el de los nacionalismos exacerbados. En la declaración final de la Conferencia se demandaba que los gobiernos nacionales reconociesen, por fin, que el gasto educativo aseguraba a las naciones una fuerza de contención y resistencia mucho más duradera que los gastos militares. Esta afirmación cobra mucho más sentido y valor si se recuerda que la mencionada declaración comienza diciendo que *"le problème est simple: il faut faire l'Europe ou il faut faire la guerre"* (Mouvement Européen, 1949, p. 18).

La historia, especialmente sus manuales, se convierte en una de las disciplinas que podemos considerar claves no sólo en cuanto a la conformación y legitimación de las identidades nacionales, sino también en lo que se refiere a la propagación de la idea de Europa. La Conferencia propone que los manuales escolares se sometan al examen de un Consejo Interuniversitario Europeo con capacidad para recomendar todas las modificaciones que estime necesarias para garantizar que estén inspirados por un espíritu europeísta y pluralista.

En realidad, la preocupación por los contenidos de los manuales de historia está presente en el panorama de la posguerra desde los inicios. La Unesco, desde su fundación, se marcó como objetivo esencial establecer una historia verdaderamente universal y con un enfoque nuevo que, en el caso de Europa, rompiese con el monopolio de su eurocentrismo. Así, celebró dos seminarios consagrados específicamente a la revisión de los manuales escolares y la historia, primero en Bruselas (1950) y luego en Sèvres (1951). En este último se ofreció un análisis de 24 manuales europeos donde se analizaban los contenidos de historia nacional, europea y universal presentes en la etapa educativa comprendida entre

los 12 y los 15 años, a través de la siguiente tabla (UNESCO, 1951a, p. 23)³:

TABLA I. Distribución espacial de los contenidos de los manuales de historia

PAÍS	H.^a NACIONAL	H.^a EUROPEA	H.^a UNIVERSAL
Suiza	67,5%	32%	1,5%
Alemania (RFA)	32%	53%	15%
Inglaterra	21%	33%	46%
Bélgica	18%	66%	16%
Escocia	51%	16%	33%
Francia	45%	55%	
Grecia	68%	16%	16%

A pesar del carácter parcialmente aleatorio de la muestra, la tabla permite alguna reflexión interesante. Alemania está en pleno período de desnazificación y sólo destina un tercio de los contenidos a la historia nacional, reservando más de la mitad a la historia europea. El caso belga es aún más claro, con dos tercios dedicados a la historia europea, si bien aquí conviene tener en cuenta que en la escuela primaria sólo se enseñaba historia nacional (Consejo de Europa, 1953). Grecia y Suiza, por su parte, representan el polo más nacionalista, con una tendencia más europea en el segundo caso. Dejamos para el final a Francia y a Inglaterra. Lo acontecido en el país galo concuerda con la cita que hemos incluido de Bayet sobre la inoportunidad de acelerar un proceso de europeización en la escuela, aunque hoy haya evolucionado en esa dirección, en la línea demandada por Sartre. Con todas las reservas de estos datos, Inglaterra presenta un interés por Europa mucho mayor que el ofrecido por estudios recientes (Gómez Carrasco y Champman, 2017).

Si la Unesco dedicó su interés a la mejora de los manuales escolares de historia como fundamento de la convivencia mundial entre las diferentes naciones y culturas, el Consejo de Europa hizo lo propio como base

³ La propia fuente indica que los datos de Inglaterra (5 manuales revisados) no son necesariamente representativos dado el gran número de manuales existentes en su sistema escolar, y que los de Francia (4 manuales) son aproximaciones, con el problema añadido de no discernir entre historia europea e historia universal. La citada fuente tampoco aclara qué se entiende por historia nacional en el caso de Inglaterra y Escocia.

de consolidación y legitimación de la idea de Europa entre las nuevas generaciones. Así, no faltan autores que hablan de una *Eurohistory* inspirada por el Movimiento Europeo e impulsada por el Consejo de Europa a partir de una serie de reuniones científicas en torno a este tema (Davies, 1996). El Consejo convocó su primer simposio en Calw, entre el 4 y el 12 de agosto de 1953, con el fin de debatir sobre la idea de Europa en la enseñanza de la historia. El simposio se dividió en dos partes. La primera se dedicó a la presentación y el debate de los diferentes informes nacionales elaborados por cada delegación acerca de los resultados de la revisión de los manuales de historia y su enseñanza en cada país, mientras que la segunda se centró en el análisis de los elementos de una concepción europea de la historia en tres momentos: la Edad Media, los tiempos modernos y la historia contemporánea. Para ello, se formularon dos cuestiones: ¿cuáles son los principios generales que concurren en la formación de una conciencia europea de la historia que conviene tener en cuenta en su enseñanza en los países europeos? y ¿cómo se deberían abordar los obstáculos a los que se enfrenta la creación de una conciencia europea común de la historia? (Consejo de Europa, 1953a). Con todo, como expresó el italiano Valsecchi, “*Le Symposium n’a pas à élaborer une nouvelle interprétation de l’histoire européenne; il doit [...] procéder à une révision politique et logique des manuels scolaires qui contribuera à faire disparaître les égoïsmes nationaux [...] il ne convient pas de rechercher dans le passé une Europe qui n’existait pas.*” (Consejo de Europa, 1953a, p. 23-24)

Hübinger, el ponente de la etapa medieval, afirma que la idea de Europa surge en los siglos XV-XVI. No hubo un medievo europeo, sino varios. No obstante, existen elementos constitutivos que serán recuperados cuando la idea de Europa tome cuerpo, como las herencias cristiana, clásica y germánica, la huella feudal, el derecho romano, la Universidad o las órdenes religiosas. Desde su punto de vista, la pervivencia de la idea imperial romana fue más un obstáculo que un elemento de unidad, ya que cada emperador pretendía ser el legítimo heredero de la tiara. No se mantuvo tampoco la unidad cultural entre occidente y oriente, como queda demostrado con la recepción de Aristóteles a través de los árabes y no de Bizancio. A pesar de la importancia de las cruzadas, no cabe hablar de una unidad europea frente al islam hasta la amenaza turca. El representante griego, Zakyrtinos, negó que el feudalismo fuese uno de los elementos forjadores de la conciencia europea, por tratarse de un

fenómeno exclusivo de la Europa central carolingia, lo cual deja fuera de esa idea de Europa a una buena parte de Italia y de España, a Escandinavia, al imperio bizantino y a los países eslavos. Tampoco aceptó la ruptura cultural entre occidente y oriente, ya que, citando a Paul Valéry, “*Partout où les noms d’Homère et de Socrate, d’Alexandre et de Platon, d’Aristote, de St. Paul ont une autorité indiscutable –là est l’Europe*” (Consejo de Europa, 1953a, p.18). En el fondo, los recursos de poder blando esgrimidos por los portavoces germano y heleno intentan legitimar una idea de Europa favorable a sus intereses regionales –occidente y oriente– en la construcción de esta nueva identidad transnacional.

En la Edad Moderna se observan ciertos principios del espíritu europeo, aparentemente contradictorios, que deben formar parte del currículo escolar. Así, la defensa de los derechos humanos es un pilar básico de la idea de Europa, con una clara secuencia entre el derecho natural –tanto católico como calvinista–, la *Bill of Rights* inglesa y la Declaración de la Revolución Francesa. Sin embargo, también es un pilar de la identidad europea la razón de estado antagónica de lo anterior, al igual que el principio dinástico y el principio democrático. La formación de Estados y la emergencia del nacionalismo no son óbice para el nacimiento simultáneo de una cierta idea de unidad europea. El universalismo es otro principio de la cultura europea que se expande y que debería abordarse en la escuela no como muestra de la hegemonía de la civilización occidental, sino de la inclusión de Europa en la historia universal, dando valor a las demás culturas.

En la Edad Contemporánea la libertad es el principio nuclear de la idea de Europa. Hubo, además, un cierto debate en torno a si la historia contemporánea y los hechos más recientes debían o no formar parte del currículo escolar. Así, se convino en que este periodo no debería ser excluido, puesto que, de lo contrario, los jóvenes accederían a él a través de la propaganda política y la prensa, sin el necesario aparato crítico que la escuela habría de proporcionarles. Finalmente, lo que sí se desechó en Calw fue presentar una recomendación al Consejo de Europa para que patrocinase la elaboración de un manual de historia general de Europa, no sólo porque sería una tarea condenada al fracaso –como había demostrado la experiencia de la Unesco–, sino también porque iría en contra de ciertas tradiciones en materia de educación, como la británica.

Conclusiones

La Europa que emerge de las ruinas de la Segunda Guerra Mundial desea –al igual que sucede a nivel global– construir unas relaciones internacionales que garanticen un futuro pacífico a las naciones. La Europa occidental, además, anhela que ese futuro pacífico se asiente en un conjunto de valores que afiancen la democracia y la libertad y alejen el fantasma de una hegemonía comunista. En este contexto, la educación tuvo un papel central para alcanzar esos objetivos. La *Primavera de Europa*, que transcurre entre el Congreso de la Haya –mayo de 1948– y el fracaso de la Comunidad Europea de Defensa –agosto de 1954–, asiste a una gran efervescencia de iniciativas y debates (ya sean auspiciados por la UO, el Movimiento Europeo, el Consejo de Europa o la Unesco) en los que la escuela está presente. Existe un denominador común: la necesidad de inculcar valores y actitudes de cooperación entre los ciudadanos de los diferentes pueblos y de mitigar, en la medida de lo posible, los estragos causados por una enseñanza sesgada por paradigmas nacionalistas y didácticas obsoletas.

En esa línea, las lenguas modernas y la historia se destacan como disciplinas merecedoras de especial atención. En el primer caso, se considera vital su conocimiento para entender mejor a los demás y fomentar el intercambio con el fin de asegurar y consolidar la cooperación internacional sin decantarse por una lengua extranjera obligatoria concreta. En el segundo, la enseñanza de la historia se caracteriza porque dos de sus puntos fuertes son la revisión de los manuales escolares –para purgarlos de mitologías nacionalistas no acordes con los nuevos tiempos de ansiada armonía universal– y el establecimiento de unas bases sólidas sobre las que cimentar los principios inspiradores y emblemáticos de la idea de Europa.

Las principales potencias europeas utilizan los numerosos encuentros sobre educación de estos años para ejercer su diplomacia cultural y poner en juego sus recursos de poder relacional, blando e inteligente. Francia y Gran Bretaña son los países que lideran la batalla por la hegemonía, no sólo en el terreno cultural, sino también en lo que respecta al diseño organizativo y relacional de la nueva Europa que está emergiendo. Francia partía con la ventaja de su experiencia con hitos como las *Écoles Françaises* de Atenas (1873) y de Roma (1875), la *Alliance Française* (1883) o la Casa Velázquez (1920). Gran Bretaña, por su parte, comienza

más tarde con hitos como el *British Council* (1934), pero sabrá aprovechar muy bien su condición de sede de los gobiernos europeos exiliados durante la Segunda Guerra Mundial para extender su poder inteligente –prestigio cultural y potencia militar victoriosa– y relacional tanto en la CAME como en el diseño organizacional del Movimiento Europeo, la UO o el Consejo de Europa. De hecho, en el ámbito de las organizaciones, se asiste a un cierto equilibrio: el Movimiento Europeo tiene dos sedes operativas (Londres y París), el Consejo de Europa es fundado en Londres pero tiene su sede en Estrasburgo, la Unesco se asienta en París pero su primer Director General es el británico Julian Huxley, los encuentros educativos promovidos por la UO se celebran primero en Gran Bretaña y luego en Francia.

En el ámbito de la hegemonía cultural, Francia puso sobre la mesa recursos de poder blando, como el prestigio de su sistema público y laico de educación, la buena reputación de sus universidades, la expansión de su lengua en las escuelas, o sus innovaciones educativas en la enseñanza de idiomas a través de la metodología directa. Por otro lado, como recursos de poder inteligente combinó el prestigio de su cultura nacional con elementos de poder duro asociados a su *grandeur*, utilizando las organizaciones intergubernamentales que hemos analizado para extender su poder relacional con la intención de convertir a la cultura francesa en matriz de la europea. Todo esto le sirvió para marcar la agenda de la Conferencia Europea de la Cultura (1949) o el seminario Unesco de Nawara Eliya, pero no para imponer, por ejemplo, su pretensión de implantar un sistema bilingüe francés-inglés en las escuelas ni la de elevar su cultura nacional al rango de líder de la cultura europea.

Por su parte, Gran Bretaña empleó recursos de poder blando como la expansión universal de su lengua –asociada a la hegemonía norteamericana–, el prestigio de sus universidades y de su sistema educativo –descentralizado y con un elevado grado de autonomía en lo que respecta, por ejemplo, a la elección de manuales– y sus innovaciones educativas en metodología audio-oral o militar –compartidas con EE. UU en este caso–, para la enseñanza de las lenguas. El poder inteligente británico se asocia con el norteamericano, que pone dinero en abundancia a través del *Network State-Private* para la difusión de la cultura anglosajona. Todo ello permite iniciar el despegue del inglés como lengua escolar extranjera predominante y rehuir la aprobación de un manual de historia europea bendecido por el Consejo de Europa en

contraposición a la tradicional autonomía de los centros a la hora de seleccionar los recursos didácticos.

La República Federal Alemana, recién nacida en 1949, también fue capaz de utilizar recursos de poder blando en la disputa por la hegemonía cultural, como el peso cuantitativo de su lengua en Europa central, el prestigio de sus instituciones universitarias y su gran legado cultural previo al nazismo. Asimismo, el poder inteligente de su centralidad geográfica resultó también eficaz, por ejemplo, en lo que se refiere al debate sobre la inclusión del feudalismo como un principio inspirador de la idea de Europa. Del mismo modo, su apuesta europeísta, muy clara en el contenido de sus manuales escolares de historia, le permite frenar las aspiraciones de quienes abogan por un sistema educativo bilingüe y ejercer un poder relacional especialmente activo en el ámbito de la diplomacia cultural.

Por lo que respecta a los países pequeños, estos también pondrán en marcha los recursos de los que disponen para participar en esta batalla cultural y educativa. Así, los países escandinavos sabrán utilizar muy bien el recurso de poder blando de su apuesta por la cooperación en el terreno educativo y de su poder relacional a través del Consejo Nórdico (1952). Grecia, por su parte, juega con el enorme prestigio de su herencia cultural para presentarse como portavoz de la periferia europea frente al modelo nórdico-carolingio a la hora de definir la idea de Europa y su transmisión en la escuela.

Finalmente, tanto la UO como el Movimiento Europeo y el Consejo de Europa van a encontrar en la escuela y en la educación un terreno fértil para legitimar sus aspiraciones de unidad. Esto constituirá, de hecho, un claro ejemplo de su poder relacional.

Referencias Bibliográficas

- Bayet, J. (1949), *Rapport a la Conférence Culturelle de l'Europe Unie sur la éducation par M. Jean Bayet*. Archives Historiques de l'Union Européenne (AHUE). Carpeta ME-540.
- Bossuat, G. (1994), *Fondateurs de l'Europe*, Paris: Belin.

- Bossuat, G. (2005), *Faire l'Europe sans defaire la France*, Bruxelles: Peter Lang.
- Consejo de Europa (1952a), *Les activites culturelles du Conseil de l'Europe. Memorandum présente au Comité des Experts culturels par le Sous-Comité Consultatif des Cinq, Strasbourg, 24-28 avril 1952*. Recuperado de <https://rm.coe.int/16804e17e7#search=%22Comite%20des%20experts%20culturels%22>, la enmienda en <http://coe.archivalware.co.uk/awweb/pdfopener?smd=1&md=1&did=938985>
- Consejo de Europa (1952b), *Asemblee Consultative. Commission des Questions culturelles et scientifiques. Proces-verbal de la séance tenue à Strasbourg, le jeudi 29 mai 1952*. Recuperado de <http://coe.archivalware.co.uk/awweb/pdfopener?smd=1&md=1&did=676729>
- Consejo de Europa (1953a), *Programme du Symposium sur la revision des manuels d'histoire, Calw (Forêt-Noire,) (4 -12 août 1953)*. Recuperado de <http://coe.archivalware.co.uk/awweb/pdfopener?smd=1&md=1&did=960396>
- Consejo de Europa (1953b), *Compte Rendu du symposium sur la révision des manuels de histoire qui s'est tenu à Calw, Fôret-Noire, du 4 au 12 de août 1953*. Recuperado de <http://coe.archivalware.co.uk/awweb/pdfopener?smd=1&md=1&did=496234>
- Davies, N. (1996), *Europe. A history*, Oxford: Oxford University Press.
- Delgado Gómez-Escalonilla, L. (2015), "Los dilemas de un propagandista reticente. El caso de Estados Unidos" en Gavari Starkie, E. y Rodríguez Jiménez, F.J. (coords.), *Estrategias de diplomacia cultural en un mundo interpolar*, pp. 47-94. Madrid: Editorial Universitaria Ramón Areces.
- Evans, I. (1948), *Carta de 19 de marzo de 1948 de Ifor Evans a Kenneth Lindsay*. AHUE. Instituto Europeo de Florencia. Carpeta ME-390.
- Flick, U. (2004), *Introducción a la investigación cualitativa*, Madrid: Morata.
- García Pedraza, R. (2015), "Desafíos a la igualdad en el siglo XXI: dificultades y retos del modelo de escuela comprensiva", *Revista Española de Educación Comparada*, 26, 135-152.
- Gavari Starkie, E. (2015), "La política educativa: ejemplo del soft-power europeo" en Gavari Starkie, E. y Rodríguez Jiménez, F.J. (coords.), *Estrategias de diplomacia cultural en un mundo interpolar*, pp. 31-46. Madrid: Editorial Universitaria Ramón Areces.

- Gómez Carrasco, C.J. y Chapman, A. (2017), “Enfoques historiográficos y representaciones sociales en los libros de texto. Un estudio comparativo, España-Francia-Inglaterra”, *Historia y memoria de la educación*, 6, 319-361.
- Haigh, A. (1974), *Cultural Diplomacy in Europe*, Strasbourg: Council of Europe.
- Judt, T., (2007), *Pasado Imperfecto. Los intelectuales franceses 1944-1956*, Madrid: Taurus.
- López Facal (2010), “Nacionalismos y europeísmos en los libros de texto: identificación e identidad nacional”, *Clio&Asociados*, 14, 9-33.
- Morales Gil, F.J., (2007), “La difusión de la metodología activa en la enseñanza de los idiomas en España: el caso del francés”, *XXI. Revista de Educación*, 9, 237-251.
- Mouvement Européen, (1949), *Conférence Européenne de la Culture. Résolutions e déclaration finale*. Lausanne: Mouvement Européen.
- Nye, J. (1990), *Bound to Lead: The Changing Nature of American Power*, New York: Basic Books.
- Nye, J. (2004), *Soft Power: the means to success in world politics*, New York: Public Affairs.
- Nye, J. (2011), *The Future of Power*, New York: Public Affairs.
- Puren, Ch. (2012) [1988], *Histoire des méthodologies de l'enseignement des langues vivantes*, Paris, Nathan-CLE International. Recuperado de https://www.aplv-languesmodernes.org/docrestreint.../puren_histoire_methodologies.pdf
- Rougemont, D. (1949), *Présentation du Rapport General par Denis de Rougemont (discours prononcé le 8 décembre 1949 à la première séance plénière)*, AHUE. Carpeta ME-538.
- Sartre, J-P. (1949), *Défense de la culture française par la culture européenne*, Paris : Centre d'Études de Politique Etrangère. AHUE. Instituto Europeo de Florencia. Carpeta ME-528.
- Scott-Smith, G. & Krabbendam, H. (eds.) (2004), *The Cultural Cold War in Western Europe, 1945-1960 (Studies in Intelligence)*, London: Routledge.
- Saunders, F. S. (1999), *Who Paid the Piper? The CIA and the Cultural Cold War*, London: Granta.
- UNESCO (1950), *Better History Textbooks*, Paris: UNESCO
- UNESCO (1951a), *International educational seminar on the teaching of history as a means of developing international understanding, Final*

report of the study group II. Sévres: UNESCO. Recuperado de <http://unesdoc.unesco.org/images/0015/001556/155685eb.pdf>

UNESCO (1951b), *International educational seminar on the teaching of history as a means of developing international understanding, Final report of the study group III*. Sévres: UNESCO. Recuperado de <http://unesdoc.unesco.org/images/0015/001556/155684eb.pdf>

UNESCO, (1954), *An account of the International seminar on the contribution of the teaching of modern languages towards education for living in a World community*, Paris: UNESCO. Recuperado de <http://unesdoc.unesco.org/images/0012/001261/126139eb.pdf>

Vigander, H. (1950), *Mutual Revision of History Textbooks in the Nordic Countries*, Paris: UNESCO

Dirección de contacto: Luís Domínguez Castro. Universidad de Vigo. Facultad de Filología y Traducción. Departamento de Historia, Arte y Geografía. Campus Lagoas-Marcosende, s/n 36210. Vigo. E-mail: dominguez@uvigo.es

La evaluación de la competencia narrativa en la educación básica¹

The assessment of narrative competence during compulsory schooling

DOI: 10.4438/1988-592X-RE-2019-383-402

María-Jesús Fernández

Manuel Montanero

Manuel Lucero

Universidad de Extremadura

Resumen

La redacción de relatos es una actividad fundamental para el desarrollo de la competencia escritora en los primeros años de escolaridad. Sin embargo, su evaluación en el contexto del aula se centra frecuentemente en aspectos superficiales, de carácter gramatical y ortográfico. Es notoria la escasez de instrumentos funcionales, con suficientes garantías de validez y fiabilidad, que faciliten la evaluación de aspectos semánticos y retóricos en la práctica del aula. En este sentido, el objetivo de esta investigación fue construir y validar una rúbrica para evaluar relatos (RER) redactados por estudiantes de entre 9 y 14 años. Para ello, se analizaron un total de 292 textos producidos por estudiantes escolarizados en estos niveles educativos. Los datos aportados apoyan la validez de la RER, lo que sugiere que tanto la estructura dimensional como la operativización de los criterios y niveles de logro son relevantes y suficientemente discriminativos para su aplicación en estos niveles educativos. También se obtuvieron datos positivos de la fiabilidad del instrumento, desde el punto de vista de su consistencia interna y de la objetividad de su aplicación por diversos

⁽¹⁾ Agradecimientos: A la Consejería de Economía, Competitividad e Innovación del Gobierno de Extremadura, y al Fondo Europeo de Desarrollo Regional por el apoyo financiero recibido (Referencia: GRU15153)

evaluadores. Los resultados del estudio-piloto muestran, sin embargo, pequeñas inconsistencias en algunos niveles de logro, como el nivel máximo del criterio “trama”. Esto se debe a la escasez de estudiantes de Secundaria que consiguen redactar tramas complejas, bien articuladas y con elementos realmente creativos. Finalmente, se discuten estos resultados y sus implicaciones para la evaluación y mejora de la competencia narrativa.

Palabras-clave: Escritura; Competencia narrativa; Relatos; Rúbrica; Educación Primaria; Educación Secundaria.

Abstract

The writing of narrative texts is considered an important activity for the development of the writing skills in the first years of schooling. However, its assessment in the classroom context is frequently based on more superficial characteristics, such as grammar and spelling. There is a notable shortage of functional and standardized writing assessment tools, offering sufficient guarantees of validity and reliability, in order to facilitate the assessment of semantic and rhetorical aspects in classroom instruction. In this sense, the present research has two core aims: (1) to develop and validate a precise and easily interpretable rubric to assess narrative compositions written by students aged between 8 and 14; and (2) to provide more information about the evolution of narrative competence throughout the compulsory education. To this effect, 292 short stories produced by students from Elementary and Secondary education were analyzed. Related to the first objective, our findings support the content-related validity of the rubric. It suggests that the rubric’s constructs or dimensions developed, the assessment categories proposed, and the levels of execution defined within each category are relevant and sufficiently discriminative to be applied in these educational stages. Also, the reliability analysis revealed positive outcomes, as much from the point of view of its internal consistency as for its inter-rater reliability, in terms of how different instructors or raters award similar scores. Regarding the second objective, the low progression of rhetorical competence during the first years of Secondary Education should be noted, as can be seen in the shortage of students who are able to write complex and well-articulated plots, with truly creative elements. Finally, results and implications for the assessment and improvement of narrative competence are discussed.

Key words: Writing; Narrative competence; Short stories; Rubric; Primary Education; Secondary Education.

Introducción

La composición de relatos es una actividad educativa fundamental para la adquisición de la competencia escritora en los primeros años de escolaridad. A partir de los 5 años los niños son ya capaces de construir verbalmente secuencias temporales sencillas, contextualizadas en marcos espacio-temporales, y que, poco a poco, van incorporando relaciones causales y motivacionales. A partir de los 7 años la mayoría de los niños son progresivamente capaces de componer por escrito dichos relatos. Los personajes se introducen por el nombre o desde la perspectiva del narrador, pero apenas se describen (Kernan, 1977); las acciones se articulan en episodios y los relatos adquieren una estructura cada vez más coherente, hasta el punto que el oyente puede predecir un final lógico durante la narración (Stadler y Ward, 2005). Las secuencias no muestran ya sólo relaciones causales entre ciertas acciones y las reacciones de los personajes, sino que se vertebran con cierta claridad en torno a un problema, una solución y en un elemento de mediación entre ambos (Iandolo, Esposito & Venuti, 2012). Entre los 8 y los 9 años se constata una mejora significativa en la cohesión y la complejidad de la estructura de la historia, que incorpora más elementos problemáticos y de mediación que retroalimentan el propio problema (Stein & Glenn, 1979). En torno a los 10 u 11 años las composiciones narrativas pueden considerarse ya relativamente maduras (Artiles & Jiménez, 2007; Yan *et al.*, 2012), con una estructura jerarquizada de episodios (primarios y secundarios), articulados en torno a un suceso inicial o un tema central (Botvin & Sutton-Smith, 1977; Ochoa-Angrino *et al.*, 2008).

Toda esta progresión se encuentra, no obstante, mediatizada por el nivel cognitivo, el dominio de la lengua y las emociones del individuo (Ryan, 2014). Así, por ejemplo, las redacciones de los niños que carecen de suficientes habilidades motoras relacionadas con la destreza escritora o que comenten muchas faltas de ortografía, manifiestan una peor organización del texto o una textualización más pobre (Berninger *et al.*, 1992; Cremin & Myhill, 2012). Cuando estos procesos se automatizan, las composiciones mejoran también en otras facetas (Graham, Harris & Fink-Chorzempa, 2002).

Desde hace décadas contamos, en definitiva, con un conocimiento relativamente amplio de la evolución de esta competencia narrativa de los niños. Sin embargo, el profesorado presta muy poca atención al

contenido semántico y a la estructura retórica de las narraciones cuando evalúa relatos (Fernández, Lucero & Montanero, 2016). La frecuencia de correcciones gramaticales y ortográficas que registran en los manuscritos de los alumnos constituye, por el contrario, el mejor predictor de la calificación global que los profesores otorgan a los textos (Lucero, Fernández & Montanero, 2018). Esta carencia podría explicar por qué los niños de Primaria mejoran significativamente estos últimos aspectos de sus borradores, pero no el contenido y la estructura retórica, cuando reciben la evaluación de sus profesores (Montanero, Fernández & Lucero, 2014).

A pesar de ser el género más frecuente en las tareas escolares de lectura y escritura, resulta también llamativo que casi ninguna de las pruebas estandarizadas más extendidas para evaluar el aprendizaje de la lectoescritura en la Educación Primaria examine específicamente la competencia retórica de los niños en tareas de composición de narraciones (véase Galve, 2005; García & Marbán 2001; Toro & Cervera, 1984; Toro, Cervera & Urío, 2002). En el contexto español el test de evaluación de procesos de escritura, PROESC (Cuetos, Ramos & Ruano, 2002) es prácticamente la única prueba estandarizada que permite evaluar específicamente dicha competencia. Su utilización en la práctica del aula es, no obstante, muy escasa, debido probablemente a su coste.

Cabe pensar que si el profesorado contara con instrumentos de evaluación fáciles de aplicar en la práctica del aula mejoraría la calidad de la retroalimentación que proporcionan a los estudiantes. En este sentido, se ha propuesto la utilización de instrumentos más sencillos, basados en escalas o rúbricas (Sundeen, 2014). Las rúbricas son un tipo particular de escalas de apreciación ordinal, en formato descriptivo, cuyos criterios aparecen operativizados en varios niveles de logro. Esta característica facilita que cada alumno conozca su situación con respecto a los objetivos que debe alcanzar en la realización de una tarea concreta (Gallavan & Kottler, 2009).

Así, por ejemplo, Benítez y García (2010) propusieron una escala de apreciación con 6 valores por criterio (desde insatisfactorio hasta excelente) que permite la evaluación de elementos retóricos de los relatos, tales como la identificación de los personajes o la construcción de la trama.

Yan *et al.* (2012) diseñaron una rúbrica de evaluación de la escritura, compuesta por 7 criterios divididos en 2 dimensiones (contenido y

organización), para alumnos de edades comprendidas entre 6 y 9 años. El instrumento valoraba la relevancia, la amplitud, el uso de conectores, la coherencia, la estructura del texto y la inteligibilidad del mismo.

Dunsmuir *et al.* (2015) propusieron otra rúbrica que permitía evaluar la composición de textos de niños con edades comprendidas entre 7 y 11 años; también contaba con 7 criterios y 4 niveles. Los aspectos que se tuvieron en cuenta, en este caso, eran la caligrafía, la ortografía, los signos de puntuación, la estructura textual, el vocabulario, la organización y la originalidad.

Se ha comprobado que este último tipo de instrumentos permiten aclarar a los estudiantes aquello que se tendrá en cuenta en la evaluación y facilitan más información que la corrección de redacciones que los profesores realizan habitualmente (Ferris & Hedgcock, 2014; Li & Lindsey, 2015). La necesidad de seleccionar uno de esos niveles en cada criterio de la rúbrica desencadena una interacción más rica en situaciones de co-evaluación, que tiene una influencia significativa en la calidad de composiciones narrativas (Montanero *et al.*, 2014; Ochoa-Angrino *et al.*, 2008), la autorregulación del aprendizaje y el sentimiento de autoeficacia (Panadero & Jonsson, 2013).

No obstante, otros autores han puesto también de manifiesto diversas debilidades, relacionadas, por un lado, con su validez y, por otro lado, con sus limitaciones de cara a la evaluación formativa de competencias (Spandel, 2006). En cuanto a la primera cuestión, Jonsson & Svingby (2007) hicieron una revisión sobre la validez y la fiabilidad de las rúbricas utilizadas en diversos estudios. Únicamente un tercio de las investigaciones tenidas en cuenta por los autores informaban de la validez de los instrumentos. Los indicadores de validez se basan principalmente en juicios de expertos, el análisis factorial confirmatorio (Baker *et al.*, 1995), o en correlaciones (superiores a 0,40) con otras medidas que se toman como criterio (Roblyer & Wiencke, 2003). En relación a la fiabilidad, las rúbricas revisadas por Jonsson & Svingby obtuvieron valores comprendidos entre 0,50 y 0,92 en la prueba Alfa de Cronbach; un grado de concordancia inter-jueces de entre el 55 y el 75%; así como índices kappa de Cohen entre 0,40 y 0,75. Tan solo uno de los trabajos revisados (Gearhart *et al.*, 1995) estaba centrado en analizar la validez de una rúbrica de evaluación de relatos.

En cuanto a la segunda cuestión, la utilización de la rúbrica en situaciones de evaluación formativa de la competencia escritora, algunos

autores han objetado el riesgo de que el instrumento encasille en cierta forma el supuesto desempeño escritor del alumnado (Fang & Wang, 2011). Diversos autores argumentan que, si bien facilita una evaluación rápida por parte del profesorado, no contribuye realmente a mejorar la capacidad de redacción del alumnado, debido a que los niveles de logro conllevan sugerencias genéricas (no personalizadas) y a menudo están formulados en un lenguaje difícil de comprender por el alumnado (Janssen *et al.*, 2015; Kohn, 2006; Panadero, Alonso-Tapia & Huertas, 2012; Schenck & Daly, 2012). Finalmente, se ha señalado la insuficiente formación de una parte del profesorado que la utiliza en el contexto del aula (East, 2009; Knoch, Read & Von Randow, 2007).

En definitiva, parece necesario generar más conocimiento sobre la validez y utilidad de este tipo de instrumentos en la práctica del aula, en particular como apoyo a la evaluación de la competencia escritora del alumnado de Educación Primaria y Secundaria. En este marco, el objetivo de este trabajo era diseñar un instrumento de evaluación de la escritura narrativa, que resulte útil y fácil de aplicar por el profesorado en la práctica del aula. En concreto, nos propusimos construir y validar una rúbrica “analítica” que permita la evaluación formativa, de manera rápida y sencilla, de textos narrativos redactados por estudiantes con edades comprendidas entre 9 y 14 años. En este sentido, es necesario aportar más evidencias de validez y fiabilidad de las que se han publicado en relación a las rúbricas anteriormente referenciadas. Analizaremos también las diferencias individuales obtenidas en la aplicación de la rúbrica a una muestra de estudiantes de diversas edades. De este modo, podemos comprobar si la rúbrica permite discriminar diferentes grados de competencia narrativa en las composiciones que redactan los estudiantes en función de su nivel educativo.

Método

Muestra

En el estudio participaron un total de 292 alumnos de 3º de Educación Primaria a 2º de Educación Secundaria Obligatoria (9-14 años). La muestra fue obtenida de 7 centros educativos de diferentes ciudades extremeñas, seleccionados por conveniencia en zonas de nivel socioeconómico medio

o medio-bajo. No participó el alumnado inmigrante con insuficiente conocimiento del español, ni los que presentaban necesidades educativas especiales. La tabla I refleja a distribución final de la muestra, en cuanto a edades, género y ciclo o nivel de escolarización.

TABLA I. Distribución de la muestra (N) según el nivel de escolarización (EP: Educación Primaria; ESO: Educación Secundaria Obligatoria), la edad (en años) y el sexo (M: masculino; F: femenino)

Nivel	Curso	Edad	Sexo	N		
I	3° EP	8-9	M	14	40	92
			F	26		
	4° EP	9-10	M	17	52	
			F	35		
II	5° EP	10-11	M	18	49	102
			F	31		
	6° EP	11-12	M	25	53	
			F	28		
III	1° ESO	12-13	M	11	25	98
			F	14		
	2° ESO	13-14	M	30	73	
			F	49		

Fuente: Elaboración propia

Instrumentos

Hoja para la redacción

A cada estudiante se le proporcionó una hoja de papel en la que debían redactar un breve relato. En la parte superior se indicaba la temática (“un niño que se convierte en el héroe de su pueblo”) y la extensión (aproximadamente una cara de un folio).

Rúbrica de evaluación de relatos (RER)

Se adaptó un instrumento de evaluación que había sido utilizado ya en investigaciones anteriores (Montanero, Lucero & Fernández, 2014), compuesto por 7 criterios de evaluación, distribuidos en dos grandes dimensiones. Por un lado, dichos criterios valoraban aspectos relativos a la organización y contenido de la narración: Marco, Tema, Trama y Creatividad. Por otro lado, se evaluaban aspectos gramaticales, léxicos y ortográficos: Oraciones, Vocabulario y Ortografía (Tabla II). Cada uno de estos criterios cuenta con 4 niveles de logro o ejecución, a los que podía asignarse las siguientes calificaciones numéricas: 0 puntos (nivel 1), 0,5 puntos (nivel 2), 1 punto (nivel 3) y 1,5 puntos (nivel 4).

PROESC

El test de Procesos de Escritura (Cuetos, Ramos & Ruano, 2002) es una prueba estandarizada, comercializada por TEA, muy extendida en la evaluación de la escritura narrativa en la educación obligatoria. Se compone de 10 criterios de evaluación agrupados en dos grandes dimensiones: Contenido y Coherencia. La primera dimensión se concreta en los siguientes criterios: Dónde y cuándo, Personajes, Suceso con consecuencias, Desenlace coherente y Creatividad. La segunda dimensión se evalúa en función de los siguientes criterios: Continuidad lógica, Sentido unitario, Figuras literarias, Oraciones complejas y Vocabulario. Para alcanzar cada uno de estos criterios se definen varios requisitos.

De acuerdo con el manual, el instrumento tiene una consistencia interna de 0,82 (coeficiente Alfa de Cronbach). Como evidencia de validez de criterio, los resultados de la aplicación de la prueba a una muestra de 839 textos arrojaron una correlación de 0,46 con las calificaciones otorgadas por un grupo de profesores. En relación a la validez estructural, los 10 criterios de los que consta la prueba fueron sometidos a un análisis de componentes con rotación varimax para obtener como resultado la división de la prueba en 3 componentes que explicaban un 58% de la varianza total.

Procedimiento

Diseño de la rúbrica de evaluación de relatos (RER)

La rúbrica, objeto de este estudio, consistió en una escala de apreciación descriptiva-ordinal (con 4 niveles por criterio, que llevaban, además asignados puntuaciones numéricas), destinada a la evaluación de relatos redactados en L1 por estudiantes de entre 9 y 14 años. Aunque inicialmente el instrumento iba destinado al alumnado escolarizado en la educación obligatoria, tras un estudio preliminar (Fernández, 2017) se concluyó que no era suficientemente válido y discriminativo para los estudiantes escolarizados en los primeros cursos de Primaria (6-8 años), ni en los últimos de Secundaria (15-16 años).

Para la selección de las variables teóricas, en las que se fundamentaron los criterios de evaluación del instrumento, se revisaron diversos antecedentes de la literatura sobre evaluación de la competencia narrativa (Benítez & García; 2010; Cuetos *et al.*, 2002; Dunsmuir *et al.*, 2015; Montanero *et al.*, 2014; Sundeen, 2014; Yan *et al.*, 2012; Thorndyke, 1977).

A partir de esta revisión bibliográfica, se identificaron 7 criterios de calidad de una composición narrativa, que resumimos a continuación. Los 4 primeros se centran en la evaluación de la organización y contenido del relato; los siguientes se centra en los aspectos gramaticales, léxicos y ortográficos.

- **Marco.** Se corresponde con la presentación de un marco espacio-temporal más o menos definido e identificable, o bien ficticio, dentro del que se suelen presentar los personajes. Lo más importante, de hecho, es la riqueza de la descripción de los personajes (desde el punto de vista físico, psicológico, etc.).
- **Tema o suceso inicial.** Es el acontecimiento o suceso inicial que desencadena el conjunto de acciones posteriores. Un buen relato alude además a los sentimientos que hacen reaccionar a los personajes y a las intenciones que guían sus acciones.
- **Trama.** Se trata de la secuencia de eventos, posteriores al suceso inicial, que dan sentido a la historia. Es muy importante la coherencia y la claridad, pero también la riqueza de los eventos y acciones que la componen.

- **Creatividad e interés.** Se relaciona con la originalidad, la creatividad y la longitud de la historia. Se valora que el relato no sea una mera reproducción de un cuento conocido, que tenga una extensión adecuada y que el autor utilice ciertos recursos estilísticos, tales como metáforas, moraleja, el sentido del humor...
- **Oraciones.** Se corresponde con la estructura de las oraciones y el uso de signos de puntuación en el texto. Un relato adecuado se caracterizaría por la inteligibilidad de sus oraciones al poseer una estructura correcta y por el uso de los signos de puntuación de manera adecuada.
- **Vocabulario.** Se refiere a la variedad y adecuación de los vocablos que el autor incluye en su relato, sin ambigüedad ni excesivas repeticiones.
- **Ortografía.** Se refiere al correcto empleo de las reglas ortográficas (en la lengua materna).

En un segundo momento, se redactó un borrador de la rúbrica en el que se definieron los 4 niveles de logro que se esperaban para cada uno de los anteriores criterios para escritores de entre 10 y 12 años.

Finalmente, la RER fue aplicada a una treintena de composiciones de estudiantes de este nivel educativo en varios ciclos. Al acabar cada ciclo, los investigadores re-definieron la redacción de algunos niveles de logro, para mejorar su claridad y objetividad, así como su capacidad discriminativa (principalmente para evitar riesgos de efecto “suelo” en algunos de los criterios). La versión final del instrumento puede verse en la tabla II.

Redacción y evaluación de relatos

La validación de la RER se basó en su aplicación a una muestra de composiciones narrativas de 292 estudiantes de 9 a 14 años.

El proceso de producción de textos se desarrolló en una sesión de una hora de duración. En una primera fase, durante unos 20 minutos, se realizó una breve introducción al concepto de narración, así como sobre las características y la estructura típica de un relato. Durante la segunda fase, cada alumno dispuso de 30 minutos para escribir una historia (inventada) cuya temática versaba sobre *un niño que se convierte en un héroe de su pueblo*, con una extensión entre una y dos caras de un folio.

Las narraciones producidas por los alumnos fueron analizadas desde tres perspectivas. Por un lado, se aplicó independientemente la rúbrica de evaluación de relatos y el PROESC. En ambos casos, cada composición recibió la puntuación correspondiente a cada criterio o nivel, únicamente cuando el texto cumplía con todos los requisitos enunciados en el mismo. En caso de faltar algún requisito, la puntuación obtenida fue la del nivel inmediatamente inferior.

Por otro lado, se seleccionaron intencionalmente cuatro maestros en activo (en centros educativos diferentes a los que se llevó a cabo el estudio) y con más de 5 años de experiencia profesional que calificaron las narraciones, utilizando una escala de 0 a 10, sin conocer las puntuaciones que dichas composiciones habían obtenido previamente con la RER y el PROESC.

Resultados

A continuación, se exponen los resultados del análisis de la validez y la fiabilidad de la rúbrica construida para evaluar composiciones narrativas.

Validez de contenido

La rúbrica fue sometida a un juicio de expertos para evaluar su validez de contenido. Se seleccionaron dos académicos, especialistas en escritura y un profesional con amplia experiencia en la Educación Primaria y Secundaria. Se les envió una hoja de validación en la que se les instaba a valorar, por un lado, la idoneidad de los criterios seleccionados y su importancia desde el punto de vista de la calidad de un relato; y, por otro lado, la claridad de la redacción de los respectivos niveles de logro.

El promedio obtenido en la valoración de la importancia (en una escala de 1 a 5) fue de 5 en todos los criterios, salvo en Creatividad (4,6), Oraciones (4,3) y Vocabulario (4,6). En cuanto a la claridad de la redacción, todos obtuvieron la máxima puntuación, salvo la trama (4,6). Los expertos hicieron, además, algunas observaciones cualitativas de carácter menor, que se tuvieron en cuenta para revisar la redacción final de la RER (tabla II).

TABLA II. Rúbrica de evaluación de relatos (RER)

Criterio	Niveles de logro	
1. Marco (Introducción)	1. No hay ninguna introducción a la historia	
	2. Sólo se nombran algunos personajes sin comentar nada de ellos.	
	3. Se cuenta cuándo y dónde ocurrió la historia y quiénes son los protagonistas pero no se describe cómo son.	
	4. Se cuenta cuándo y dónde ocurrió la historia. Se describe quiénes son los protagonistas y cómo son (físicamente, personalidad, etc.).	
2. Tema o suceso inicial	1. No se explica qué pasó al principio	
	2. Se explica algo que pasó al principio, pero no se entiende bien.	
	3. Se explica un poco lo que pasó al principio y cómo se sintieron los personajes.	
	4. Se describe con claridad todo lo que pasó al principio, cómo se sintieron los personajes y qué se propusieron hacer.	
3. Trama (episodios y resolución)	1. En la historia no pasa nada o no se entiende nada de lo que pasa.	
	2. La historia es un poco liosa o algunos hechos no tienen mucho sentido.	
	3. La historia es sencilla. Se explica con suficiente claridad lo que pasó después y al final.	
	4. Aunque la historia es larga, se explica muy bien todo lo que pasó después (en varios momentos bien ordenados), y cómo quedaron al final los personajes.	
4. Creatividad e interés	1. La historia está copiada o no se entiende.	
	2. La historia es original y se entiende, pero es demasiado corta o muy aburrida	
	3. La historia es original y entretenida	
	4. La historia es original y entretenida. Es bonita, tiene sentido del humor o nos enseña algo	
5. Oraciones	1. La mayoría de las frases no se entienden porque están mal construidas.	
	2. La mayoría de las frases se entienden bien, aunque algunas están mal hechas.	
	3. Las frases se entienden bien, pero hay pocos signos de puntuación (o sólo son comas)	
	4. Las frases se entienden bien y la mayoría están bien puntuadas, incluyendo los diálogos.	
6. Vocabulario	1. El vocabulario es muy pobre y con muchos errores (algunas palabras no son las más adecuadas para lo que se quiere decir).	
	2. Las palabras son correctas pero se repiten mucho.	
	3. Se utiliza un vocabulario correcto y variado (que no se repite demasiado).	
	4. Se utiliza un vocabulario rico, con algunas palabras poco conocidas.	
7. Ortografía	1. Hay muchas faltas de ortografía (una o más en cada oración).	
	2. Hay bastantes faltas de ortografía (una por cada dos oraciones aproximadamente)	
	3. Hay pocas faltas de ortografía	
	4. No hay ninguna falta de ortografía	

Fuente: adaptada de Montanero, Lucero y Fernández (2014)

Validez de constructo: análisis de las diferencias individuales en función del nivel educativo

La versión final de la RER se aplicó a una muestra de 292 estudiantes de Educación Primaria y Secundaria. La tabla III refleja el promedio de los resultados de cada criterio en los tres niveles educativos que se compararon. Entre el primer y segundo nivel hay diferencias significativas en cuanto al marco ($t=2,48$; $p<0,05$), al tema ($t=3,07$; $p<0,01$), las oraciones ($t=2,69$; $p<0,01$), la ortografía ($t=4,48$; $p<0,01$); así como en cuanto al subtotal de la dimensión Organización y contenido ($t = 3,18$; $p<0,01$), al subtotal de la dimensión Aspectos gramaticales, léxicos y ortográficos ($t = 4,25$; $p<0,01$) y al total de la RER ($t=4,42$; $p<0,01$). Entre el segundo y el tercer nivel se detectan diferencias significativas, en cuanto a la creatividad ($t=3,52$; $p<0,01$) y la ortografía ($t=2,58$; $p<0,05$); así como al subtotal de Organización y contenido ($t=3,18$; $p<0,01$), al subtotal de Aspectos gramaticales ($t=2,58$; $p<0,05$) y al total de la rúbrica ($t=3,50$; $p<0,01$).

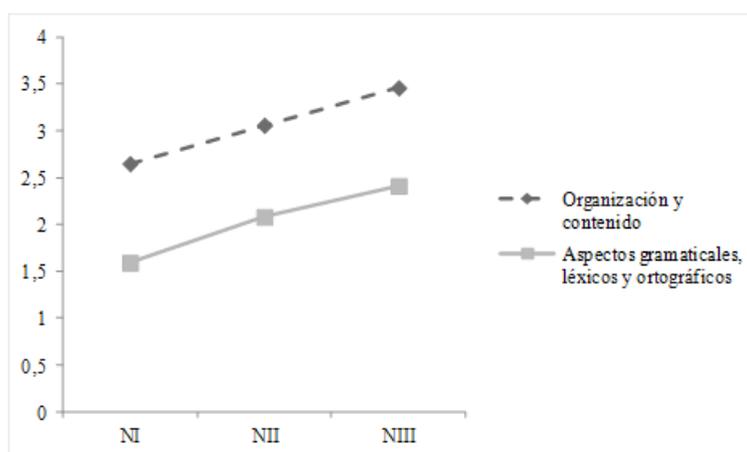
TABLA III. Diferencias en los resultados de la rúbrica de evaluación de relatos (RER), en función del nivel de escolarización (NI: 3º y 4º de Educación Primaria; NII: 5º y 6º de Educación Primaria; NIII: 1º y 2º de Educación Secundaria Obligatoria)

RER	Media (M) y desviación típica (DT)						Diferencia de medias (DM)		
	Nivel I		Nivel II		Nivel III		NI-II	NI-III	NII-III
	M	DT	M	DT	M	DT	DM	DM	DM
Marco	0,79	0,4	0,92	0,35	0,99	0,38	-0,13*	-0,2**	-0,07
Tema	0,63	0,31	0,77	0,34	0,87	0,37	-0,14*	-0,24**	-0,1
Trama	0,77	0,32	0,83	0,37	0,89	0,28	-0,06	-0,12*	-0,06
Creatividad	0,47	0,32	0,54	0,38	0,7	0,28	-0,07	-0,23**	-0,16*
Subtotal Organización y contenido	2,65	0,87	3,06	0,95	3,46	0,8	-0,41**	-0,81**	-0,4**
Oraciones	0,67	0,33	0,81	0,39	0,91	0,41	-0,14*	-0,24**	-0,1
Vocabulario	0,63	0,35	0,71	0,35	0,8	0,45	-0,08	-0,17**	-0,09
Ortografía	0,28	0,45	0,55	0,40	0,7	0,4	-0,27**	-0,42**	-0,15*
Subtotal Aspectos gramaticales, léxicos y ortográficos	1,59	0,75	2,08	0,85	2,41	0,95	-0,49**	-0,82**	-0,33*
Total	4,23	1,35	5,14	1,50	5,87	1,43	-0,91**	-1,64**	-0,73**

Fuente: Elaboración propia. (*) $p<0,05$; (**) $p<0,01$

Un indicador de la validez de un instrumento es su capacidad de discriminar entre diferentes niveles de desarrollo de la competencia que supuestamente evalúa. En el siguiente gráfico se observa una relación aumento progresivo y lineal de los promedios de la rúbrica a medida que aumenta el nivel en el que están escolarizados los estudiantes que redactaron los textos.

GRÁFICO I. Evolución de los criterios globales en función del nivel de escolarización (NI: 3º y 4º de Educación Primaria; NII: 5º y 6º de Educación Primaria; NIII: 1º y 2º de Educación Secundaria Obligatoria).



Fuente: Elaboración propia

Más específicamente, la tabla IV recoge la distribución de los sujetos en cada uno de los niveles de logro de los criterios de la RER.

TABLA IV. Porcentaje de estudiantes localizados en cada nivel de logro (NL) de la rúbrica de evaluación de relatos (RER) en función del nivel de escolarización (NI: 3° y 4° de Educación Primaria; NII: 5° y 6° de Educación Primaria; NIII: 1° y 2° de Educación Secundaria Obligatoria)

Criterio	NL	NI	NII	NIII
1. Marco	1	1,1	1	0
	2	58,7	30,4	28,6
	3	21,7	52	43,9
	4	18,5	16,7	27,6
2. Tema o suceso inicial	1	5,4	1	0
	2	68,5	53,9	42,9
	3	21,7	35,3	39,8
	4	4,3	9,8	17,3
3. Trama	1	2,2	4,9	0
	2	47,8	35,3	29,6
	3	44,6	48	63,3
	4	5,4	11,8	7,1
4. Creatividad	1	22,8	21,6	3,1
	2	62	52	53,1
	3	14,1	23,5	43,9
	4	1,1	2,9	0
5. Oraciones	1	7,6	6,9	2
	2	52,2	35,3	38,8
	3	38	46,1	34,7
	4	2,2	11,8	24,5
6. Vocabulario	1	12	5,9	10,2
	2	52,2	52	36,7
	3	33,7	36,3	35,7
	4	2,2	5,9	17,3
7. Ortografía	1	68,4	25,5	15,3
	2	8,7	39,2	33,7
	3	20,7	34,3	46,9
	4	2,2	1	4,1

Fuente: Elaboración propia

Como cabía esperar, el *contenido* de los relatos se va haciendo más complejo a medida que el alumno va evolucionando por los niveles del sistema educativo. Así, por ejemplo, el alumnado de tercer curso no suele referir al marco *espacio-temporal* en el que se desarrolla el relato. En los últimos cursos de Primaria, por el contrario, se menciona frecuentemente el nombre del lugar y se utilizan fórmulas específicas del género para aludir al contexto temporal (por ejemplo, “érase una vez” o “había una vez”). En Primaria en general la narración gira casi siempre alrededor de un único personaje, que apenas se describe o sobre el que se señala alguna característica física. En ocasiones, se alude a personajes que se desconocía que existían o cuyo papel en la trama no queda del todo claro. En Secundaria, en cambio, la historia se contextualiza y se describe a los personajes con mayor detalle, incluyendo rasgos psicológicos o psicosociales (“Luis no era buen estudiante, pero se caracterizaba por ser solidario y tranquilo”). Además, aparecen casi siempre *personajes secundarios* que participan en eventos relevantes.

La RER pone también de manifiesto cómo la *trama* se va haciendo más compleja a lo largo de los niveles educativos. En bastantes textos de 3º de Educación Primaria prácticamente no ocurre nada (son meras descripciones sin sucesos problemáticos que desencadenen acciones posteriores); mientras que en otros la trama resulta muy confusa, con disrupciones en el hilo conductor y desenlaces poco coherentes con el inicio de la historia. En este nivel muchos niños no componen una historia original, sino que adaptan argumentos de cuentos conocidos. A partir de 5º curso se aprecia una mejora significativa en la coherencia. Se alude ya explícitamente a los sentimientos de los personajes y lo que éstos se propusieron hacer una vez delimitado el problema inicial, aunque todavía en bastantes textos las acciones se yuxtaponen o carecen de una lógica causal o motivacional. Los relatos son más extensos y creativos e incorporan frecuentemente diálogos y moralejas. En los cursos de Secundaria la trama se extiende, al incorporar más episodios y al aparecer, en ocasiones, temas multi-problemáticos que generan ciclos de desenlace.

Sin embargo, sorprendentemente el porcentaje de estudiantes que alcanzan el máximo nivel de logro en la trama es ligeramente inferior a los que lo consiguen en los últimos cursos de Primaria. Los textos de Secundaria son más extensos, pero no mucho más creativos que

en el nivel anterior y curiosamente suelen carecer de diálogos y otros elementos que enriquecen el relato.

En cuanto a los aspectos *formales y gramaticales*, los resultados de la RER son también coherentes con la evolución que se esperaba en función del nivel educativo del alumnado. En los textos de 3º y 4º predominan las oraciones simples y las compuestas coordinadas de un modo muy pobre (con un abuso de la conjunción “y”, así como de anáforas fieles, que resultan excesivamente repetitivas). Un 70% de los estudiantes de 3º y 4º tienen más de 10 faltas en el texto (sin contar tildes); mientras que en 5º y 6º la mayoría tiene más de 5. Hasta Secundaria prácticamente no aparecen oraciones subordinadas sustantivas y adjetivas y un empleo algo más preciso de los signos de puntuación (aunque el abuso de la coma, donde debería aparecer un punto o un punto y coma, es una constante). El vocabulario y los recursos anafóricos, no obstante, siguen resultando poco variados repetitivos en la mayoría de los alumnos de esta etapa. Casi todos los textos de Secundaria presentan todavía errores ortográficos, aunque la mayoría contienen menos de 5 faltas.

Normalidad

En la siguiente tabla se observa que las puntuaciones obtenidas en las composiciones de los diferentes niveles educativos en cada una de las dimensiones se adecuan suficientemente a los modelos de una distribución normal, teniendo en cuenta la bondad de ajuste, la asimetría y la curtosis obtenida en cada caso.

TABLA V. Distribuciones de las puntuaciones de la rúbrica de evaluación de relatos (RER) por dimensiones en cada nivel de escolarización (NI: 3° y 4° de Educación Primaria; NII: 5° y 6° de Educación Primaria; NIII: 1° y 2° de Educación Secundaria Obligatoria).

Dimensiones	Nivel I	Nivel II	Nivel III
<i>Organización y contenido</i>	Media: 2,65 DT: 0,87 Asimetría: 0,13 Curtosis: -0,20 Z de K-S: 1,27	Media: 3,06 DT: 0,95 Asimetría: 0,14 Curtosis: -0,59 Z de K-S: 1,48	Media: 3,46 DT: 0,80 Asimetría: 0,06 Curtosis: -0,67 Z de K-S: 1,43
<i>Aspectos gramaticales, léxicos y ortográficos</i>	Media: 1,59 DT: 0,75 Asimetría: 0,22 Curtosis: 0,01 Z de K-S: 1,49	Media: 2,08 DT: 0,85 Asimetría: 0,10 Curtosis: 0,03 Z de K-S: 1,36	Media: 2,41 DT: 0,95 Asimetría: 0,05 Curtosis: -0,51 Z de K-S: 1,24

Fuente: Elaboración propia

En las gráficas II, III, IV se puede observar un mejor ajuste en los niveles educativos I y II. En la dimensión *Organización y contenido* la curva resulta en todos los niveles ligeramente platicúrtica, especialmente en el nivel III.

GRÁFICO II. Distribución de las puntuaciones en Nivel I (3° y 4° de Primaria)

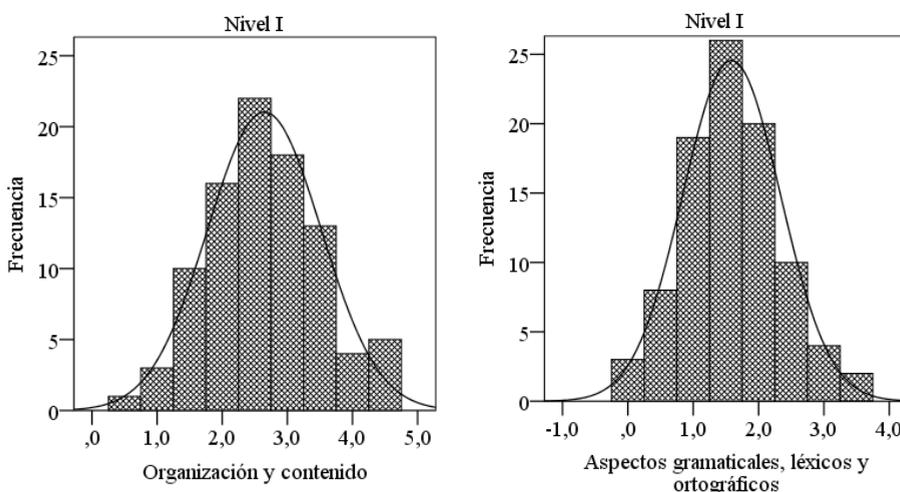


GRÁFICO III. Distribución de las puntuaciones en Nivel II (5° y 6° de Primaria)

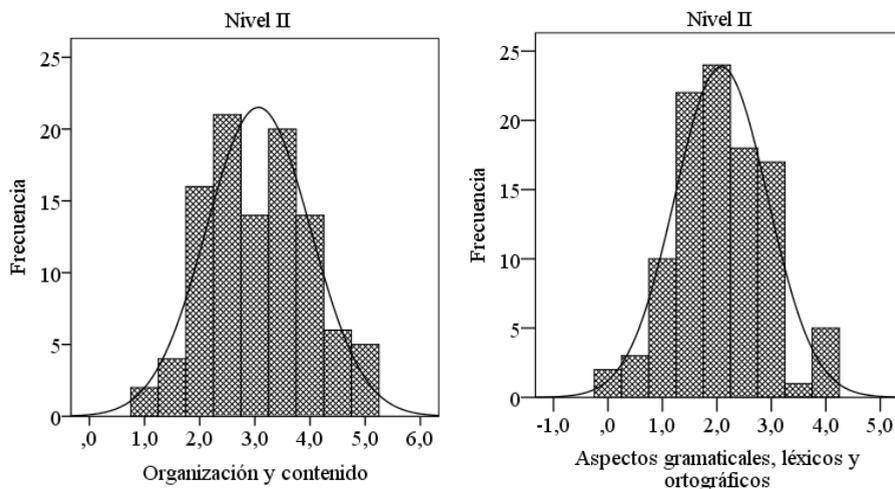
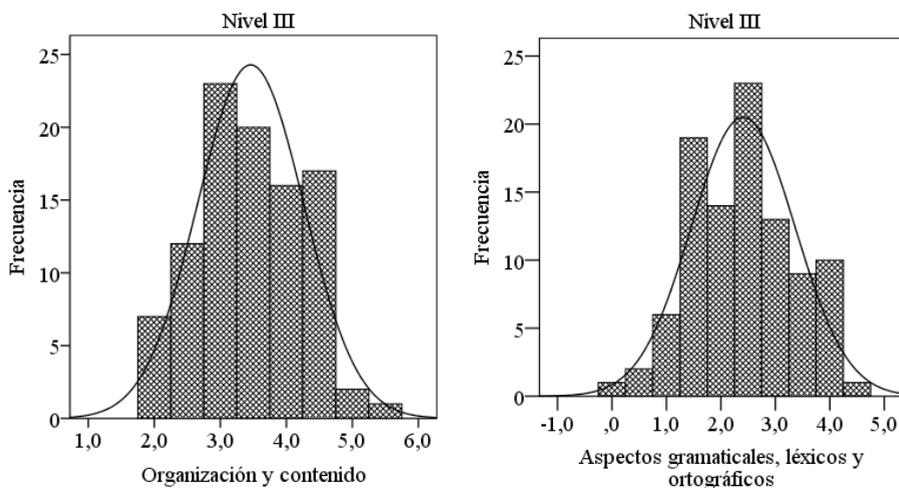


GRÁFICO IV. Distribución de las puntuaciones en Nivel III (1° y 2° de ESO)



Validez estructural

Para valorar la estructura de las dimensiones y criterio de la RER se realizó un análisis factorial confirmatorio de componentes principales de la rúbrica. Previamente se aplicaron dos pruebas que evidenciaron que el conjunto de datos era adecuado para ser factorizado: la medida de la adecuación muestral de Kaiser-Meyer-Olkin (0,75) y la prueba de esfericidad de Barlett (Chi-cuadrado=161,45; $p < 0,01$).

En la tabla VI se presenta la matriz factorial de componentes rotados, según el procedimiento varimax, en la que aparecen los 2 componentes extraídos y la saturación de las variables en cada uno de ellos.

TABLA VI. Matriz de componentes rotados según varimax

Criterios	Componentes	
	1	2
Marco	0,206	0,421
Tema	0,715	0,064
Trama	0,711	0,174
Creatividad	0,739	0,200
Oraciones	0,356	0,600
Vocabulario	0,012	0,755
Ortografía	0,083	0,697

Fuente: Elaboración propia

Se observan en negrita los valores de cada criterio que se corresponden con una saturación mayor en cada uno de los componentes. Para vincular un factor con un componente se aconseja que el primero de ellos tenga un peso factorial significativo (superior a 0,50). Aplicando un criterio más laxo (superior a 0,2) el criterio Marco podría considerarse realmente en cualquiera de los dos componentes. Se ha incluido en el primero porque resulta más coherente con su validez teórica de constructo (a pesar de contar con un peso factorial algo menor).

Los 7 factores quedan, por tanto, agrupados en 2 componentes que en su conjunto explican el 49% de la varianza total, teniendo mayor peso el primer componente que explica un 34%. De acuerdo con la estructura

del instrumento, los componentes parecen por tanto coherentes con las dimensiones Contenido y organización (componente 1), así como con la dimensión relativa a los aspectos gramaticales, léxicos y ortográficos (Componente 2).

Validez de criterio

La correlación entre el sumatorio de las puntuaciones obtenidas en la RER y en el test estandarizado PROESC, a partir de la evaluación de cada uno de los textos, resultó muy elevada $r=0,79$ ($p<0,01$), aunque no tanto con las calificaciones que el profesorado en activo en estos mismos niveles educativos otorgó a los textos ($r=0,41$; $p<0,01$). No obstante, esta última correlación fue superior a la que obtuvimos entre dichas calificaciones y los resultados del PROESC ($r=0,30$; $p<0,01$).

Más específicamente, en la tabla VII se muestra la correlación con cada uno de los criterios del PROESC que evalúan aspectos similares (no se ha incluido el criterio “ortografía” de la rúbrica por que no existe ninguno similar en el PROESC).

TABLA VII. Correlación de Pearson (r) entre los criterios específicos de los dos instrumentos

RER	PROESC	r
Marco	Dónde y cuándo Personajes	0,8* 0,27*
Tema o suceso inicial	Suceso con consecuencias	0,28*
Trama	Suceso con consecuencias Desenlace coherente Continuidad lógica Sentido unitario	0,29* 0,57* 0,43* 0,46*
Creatividad	Creatividad Figuras literarias	0,34* 0,14
Oraciones	Oraciones complejas	0,28*
Vocabulario	Vocabulario	0,48*
Ortografía	-	-

Nota: Correlaciones estadísticamente significativas: (*) $p < 0,01$.

Fuente: Elaboración propia

Consistencia interna

Al analizar la homogeneidad de los 4 elementos de los que consta la dimensión *Organización y contenido* del instrumento se obtuvo un Alfa de Cronbach de 0,57, mientras que en la dimensión *Aspectos gramaticales, léxicos y ortográficos* fue de 0,60. Se trata de un valor satisfactorio y superior al que hemos obtenido con el PROESC (0,37 en ambas dimensiones). Se observa también que si se eliminase algún criterio de la rúbrica el valor de Alfa no mejoraría.

Fiabilidad interjueces

Para calcular la fiabilidad en la aplicación de la RER se analizó el grado de acuerdo entre 3 evaluadores sobre una muestra de 30 narraciones, escogidas al azar, tras varias sesiones de entrenamiento en su aplicación. Los índices de fiabilidad obtenidos por el método kappa de Cohen, fueron superiores a 0,85 ($p < 0,01$) en todos los criterios. Nuevamente, los resultados fueron ligeramente superiores a los obtenidos con el PROESC.

Conclusiones

La evaluación formativa de la escritura es una práctica habitual del aula que se encuentra mediatizada por las opiniones subjetivas de los maestros sobre los aspectos que reflejan una buena escritura (Barton, Hamilton, & Ivancic, 2000). La revisión de la bibliografía que hemos realizado en la introducción de este trabajo muestra la escasez de instrumentos objetivos, con suficientes garantías de validez y fiabilidad, que sean accesibles para el profesorado y permitan, al mismo tiempo una utilización funcional en el aula. Las escalas ordinales de apreciación en formato descriptivo (o rúbricas) se adecuan muy bien a ese objetivo: por un lado, por que ofrecen indicaciones concretas sobre el nivel de logro alcanzado por el estudiante en los criterios específicos de evaluación de una competencia o tarea, así como las siguientes mejoras o avances que puede realizar en la progresión de su aprendizaje; por otro lado, porque su diseño funcional e intuitivo (normalmente en una sola hoja), facilita

su utilización en la práctica del aula, no sólo por el profesorado, sino también por los propios estudiantes.

En este sentido, el principal objetivo de este trabajo era diseñar y validar una rúbrica para la evaluación de relatos (RER) en los primeros años de la escolaridad obligatoria.

Los datos aportados en primer lugar en el apartado de Resultados apoyan la validez de contenido de la rúbrica, lo que sugiere que tanto la estructura dimensional como la operativización de los criterios y niveles de logro son relevantes y representativos del constructo para su evaluación en los tres niveles escolares seleccionados. Los expertos (investigadores y profesionales de la educación) que han tenido la oportunidad de revisar y utilizar la RER han valorado muy positivamente, tanto su construcción como su utilidad práctica. El análisis factorial de los datos obtenidos tras su aplicación a una amplia muestra de relatos ha confirmado la adecuación de la estructura de dos factores o dimensiones que agrupan los diversos criterios de evaluación. Al aplicar la rúbrica a la evaluación de relatos redactados por estudiantes de diferentes edades, se ha comprobado también que el instrumento discrimina aceptablemente la progresiva evolución en la madurez en la competencia narrativa que se espera en alumnos de 9 a 14 años. Además, la correlación global de 0,79, con una prueba estandarizada (el PROESC), puede considerarse más que suficiente como indicador de *validez criterial* (Roblyer & Wiencke, 2003). Otro importante indicador de validez es la correlación con las calificaciones que el profesorado en activo en estos mismos niveles educativos otorgó a los textos (Gearhart *et al.*, 1995), que resultó incluso superior a la que hemos registrado con el PROESC.

En segundo lugar, hemos obtenido también datos positivos de la fiabilidad del instrumento, desde el punto de vista de su consistencia interna y de la objetividad de su aplicación por diversos evaluadores. En cuanto a la consistencia (medida con la prueba Alfa de Cronbach), los resultados se encuentran comprendidos dentro de los parámetros que diversos expertos recomiendan en su revisión de investigaciones sobre rúbricas (Graham & Perin, 2007; Jonsson & Svingby, 2007; Knoch, 2009). Los resultados de la aplicación del índice kappa de Cohen a nuestra rúbrica (superiores a 0,85 en todos los criterios) son también mejores que en la mayoría de dichos trabajos.

En definitiva, la rúbrica estudiada parece un instrumento sencillo, válido y fiable para la evaluación de relatos redactados por estudiantes

de 9 a 14 años. Dadas sus características y nuestra experiencia en trabajos previos (Montanero *et al.*, 2014), cabe suponer además su potencialidad para apoyar actividades de evaluación y co-evaluación formativa en el contexto del aula.

Por último, las diferencias encontradas entre la calidad de los relatos en diferentes niveles educativos es, en general, coherente con lo que se esperaba. Cabe señalar, sin embargo, la escasez de estudiantes de los últimos cursos que consiguen redactar tramas complejas, bien articuladas y con elementos realmente creativos. En particular, se observa una débil progresión en cuanto a la competencia retórica de la mayoría del alumnado de los primeros cursos de Secundaria, a la que ya aludían otras investigaciones previas en nuestro sistema educativo (Artiles & Jiménez, 2007; Salvador, 2000). Una posible explicación para este resultado es que la RER responde a una estructura clásica de relato sencillo y lineal. No identifica una diversidad de recursos (como las analepsias, el contraste de personajes en diferentes tiempos, etc.) que escritores más avanzados pueden utilizar para enriquecer el relato; lo que limita su capacidad discriminativa en un rango tan amplio de edad. En futuros estudios pretendemos adecuar y validar diferentes versiones de la RER para rangos de edad de sólo 2 años.

En cualquier caso, la incorporación al aula de instrumentos que faciliten una evaluación formativa de esta competencia puede contribuir a potenciarla. Para confirmar la utilidad de este instrumento en futuros trabajos pretendemos también analizar longitudinalmente la evolución de la calidad de los relatos de estudiantes que participen en actividades de evaluación con esta rúbrica, en comparación con otros que no lo hagan.

Referencias bibliográficas

Artiles, C., & Jiménez, J. E. (2007). *Proescri Primaria. Prueba de evaluación de procesos cognitivos en la escritura*. Canarias: Consejería de Educación, Cultura y Deportes del Gobierno de Canarias. Dirección General de Ordenación e Innovación Educativa.

- Baker, E.L., Abedi, J.; Linn, R. L., & Niemi, D. (1995). Dimensionality and generalizability of domain-independent performance assessments. *Journal of Educational Research* 89 (4), 197–205.
- Barton, D., Hamilton, M., & Ivanic, R. (2000). *Situated Literacies: reading and writing in context*. Routledge: London.
- Benítez, R., & García, G. (2010). El razonamiento analógico verbal: una habilidad cognitiva esencial de la producción escrita. *Onomázein*, 22 (2), 165-194.-
- Berninger, V.W, Yates, C., Cartwright, A., Rutberg, J., Remy, E., & Abbott, R. (1992). Lower-level developmental skills in beginning writing. *Reading and Writing: An interdisciplinary Journal*, 4, 257-280.
- Botvin, G. J., & Sutton-Smith, B. (1977). The development of structural complexity in children's fantasy narratives. *Developmental Psychology*, 13 (4), 377-388.
- Cremin, T., & Myhill, D. (2012). *Writing Voices: Creating communities of writers*. London: Routledge.
- Cuetos, F., Ramos, J. L., & Ruano, E. (2002). *PROESC: Bateria de evaluación de los procesos de escritura*. Madrid: TEA Ediciones.
- Dunsmuir, S. Kyriacou, M. Batuwitige, S. Hinson, E. Ingram, V., & O'Sullivan, S. (2015). An evaluation of the Writing Assessment Measure (WAM) for children's narrative writing. *Assessing Writing*, 23 (2015) 1–18.
- East, M. (2009). Evaluating the reliability of a detailed analytic scoring rubric for foreign language writing. *Assessing Writing*, 14, 88-115. DOI: 10.1016/j.asw.2009.04.001
- Fang, Z., & Wang, Z. (2011). Beyond rubrics: Using functional language analysis to evaluate students writing. *Australian Journal of Language and Literacy*, 34 (2), 147-165.
- Fernández, M. J., Lucero, M., & Montanero, M. (2016). Rojo sobre negro. ¿Cómo evalúan los maestros las redacciones de sus estudiantes? *Revista de educación*, 372 (abril-junio), 63-82.
- Fernández, M. J. (2017). Co-evaluación de textos narrativos en la Educación Obligatoria. Universidad de Extremadura, Badajoz, España.
- Ferris, D. R., & Hedgcock, J. (2014). *Teaching ESL composition: Purposes, process, and practice* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum.
- Gallavan, N., G. & Kottler, E. (2009). Constructing rubrics and assessing progress collaboratively with social studies students, social studies students. *The Social Studies*, 100 (4), 154-158.

- Galve, J. (2005). BECOLE. *Batería de Evaluación Cognitiva de la Lectura y Escritura*. Instituto de Orientación Psicológica. Madrid: Editorial EOS.
- García, J. N., & Marbán, J. M. (2001): Instrumento de Evaluación Individual de los Procesos Cognitivos de la Escritura (IEPCE). En J. N. García, *Dificultades de aprendizaje e intervención psicopedagógica* (pp. 141-149). Barcelona: Ariel.
- Gearhart, M., Herman, J. L., Novak, J. R., & Wolf, S. A. (1995). Toward the instructional utility of large-scale writing assessment: Validation of a new narrative rubric. *Assessing Writing*, 2, 207–242.
- Graham, S., & Perin, D. (2007). A meta-analysis of writing instruction for adolescent students. *Journal of Educational Psychology*, 99 (3), 445–476.
- Graham, S., Harris, K.R., & Fink-Chorzempa, B. (2002). Contributions of spelling instruction to the spelling, writing, and reading of poor spellers. *Journal of Educational Psychology*, 94, 669-686.
- Iandolo, G., Esposito, G., & Venuti, P. (2012). Cohesión, micro-organización, estructura narrativa y competencias verbales entre tres y once años: el desarrollo narrativo formal. *Estudios de Psicología*, 34(2), 141-160.
- Janssen, G., Meier, V., & Trace, J. (2015). Building a better rubric: Mixed methods rubric revision. *Assessing Writing*, 26, 51-66. DOI: 10.1016/j.asw.2015.07.002
- Jonsson A., & Svingby, G. (2007). The Use Of Scoring Rubrics: Reliability, Validity And Educational Consequences, *Educational Research Review*, 2, 130–144. DOI: 10.1016/j.edurev.2007.05.002
- Kernan, Keith T. (1977). Semantic and expressive elaboration in children's narratives. *Child discourse*, 91-102.
- Knoch, U. (2009). Diagnostic assessment of writing: A comparison of two rating scales. *Language Testing*, 26(2), 275-304. DOI: 10.1177/0265532208101008
- Knoch, U., Read, J., & Von Randow, J. (2007). Re-training writing raters online: How does it compare with face-to-face training? *Assessing Writing*, 12(1), 26–43. DOI: 10.1016/j.asw.2007.04.001
- Kohn, A. (2006). The trouble with rubrics. *English Journal*, 95(4), 12–15.
- Li, J., & Lindsey, P. (2015). Understanding variations between student and teacher application of rubrics. *Assessing Writing*, 26, 67–79. DOI: 10.1016/j.asw.2015.07.003

- Lucero, M., Fernández, M. J., & Montanero, M (2018). Assessment of short stories in Elementary and Secondary Education. A study of teachers' responses to students' texts. *Studies in Educational Evaluation*, 59, 158-167.
- Montanero, M., Lucero, M., & Fernández, M. J. (2014). Iterative co-evaluation with a rubric of narrative texts in Primary Education. *Infancia y Aprendizaje: Journal for the Study of Education and Development*, 37 (1), 184-220. DOI: 10.1080/02103702.2014.881653
- Ochoa-Angrino, S., Aragón, L., Correa, M., & Mosquera, S. (2008). Funcionamiento metacognitivo de niños escolares en la escritura de un texto narrativo antes y después de una pauta de corrección conjunta [Versión electrónica], *Acta Colombiana de Psicología*, 11 (2), 77-88.
- Panadero, E., & Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational Research Review*, 9(0), 129-144. DOI: <http://dx.doi.org/10.1016/j.edurev.2013.01.002>
- Panadero, E., Alonso-Tapia, J., & Huertas, J. A. (2012). Rubrics and self assessment scripts effects on self-regulation, learning and self-efficacy in secondary education. *Learning and Individual Differences* (0). DOI: 10.1016/j.lindif. 2012.04.007
- Roblyer, M. D., & Wiencke, W. R. (2003). Design and use of a rubric to assess and encourage interactive qualities in distance courses. *American Journal of Distance Education*, 17, 77-99.
- Ryan, M. (2014). Reflexive writers; Re-thinking writing development and assessment in schools. *Assessing Writing*, 22, 60-74. DOI:10.1016/j.asw.2014.08.002
- Salvador, F. (2000). Habilidades narrativas de alumnos de Educación Primaria en la producción de textos escritos. *Enseñanza*, 17-18, 145-163.
- Schenck, A. D., & Daly, E. (2012). Building a better mousetrap: Replacing subjective writing rubrics with more empirically-sound alternatives for EFL learners. *Creative Education*, 3 (8), 1310-1325. DOI: 10.4236/ce.2012.38193
- Spandel, V. (2006). Speaking my mind: In defense of rubrics. *English Journal*, 96 (1), 19-22.
- Stadler, M. A., & Ward, G. C. (2005). Supporting the narrative development of young children. *Early Childhood Education Journal*, 33, 2, 73-80.

- Stein, N., & Glenn, C. (1979). An analysis of story comprehension in elementary school children. En R. O. Freedle (Ed.), *Advances in discourse processing*. Vol. 2. Norwood, Nueva Jersey: Ablex.
- Sundeen, T. H. (2014). Instructional rubrics: Effects of presentation options on writing quality. *Assessing Writing*, 21, 74-88. DOI: 10.1016/j.asw.2014.03.003
- Thorndyke, P. W. (1977). Cognitive structures in comprehension and memory of narrative discourse. *Cognitive Psychology*, 9, 77-110.
- Toro, J., & Cervera, M. (1984). *TALE. Test de análisis de la lectoescritura*. Madrid: Aprendizaje Visor.
- Toro, J., Cervera, M., & Urío, C. (2002). *Escalas Magallanes de Lectura y Escritura. TALE-2000*. Vizcaya: Albor-COHS.
- Yan, C. M. W., McBride-Chang, C., Wagner, R. K., Zhang, J., Wong, A. M. Y., & Shu, H. (2012). Writing quality in Chinese children: speed and fluency matter. *Reading and Writing*, 25 (7), 1499-1521.

Dirección de contacto: María Jesús Fernández Sánchez, Universidad de Extremadura, Facultad de Educación, Departamento de Ciencias de la Educación, Avda. de Elvas s/n. 06006 Badajoz, España. E-mail: mafernandezs@unex.es

Las revistas de educación a nivel mundial: un análisis de las publicaciones incluidas en el Journal Citation Reports (JCR) del 2016

Education journals worldwide: an analysis of the publications included in the 2016 Journal Citation Reports (JCR)

DOI: 10.4438/1988-592X-RE-2019-383-403

Julia Haba-Osca
Francisco González-Sala

Universitat de València

Julia Osca-Lluch

INGENIO (CSIC-UPV)

Resumen

Introducción: La evaluación de la actividad investigadora en España, otorga reconocimiento personal y social a profesores e investigadores y condiciona su carrera académica. Los artículos publicados en revistas incluidas en el Journal Citation Reports (JCR), son los más necesarios para conseguir la acreditación personal y sexenios de investigación. En este trabajo, se analiza la situación de las revistas de educación incluidas actualmente en los JCR (SCIE y SSCI). **Metodología:** la muestra la conformaron 340 revistas incluidas en las categorías temáticas *Education & Educational Research*, *Special*, *Education* y *Psychology Educational* de la edición del JCR del Science Citation Index Expanded (SCIE) del 2016. Se estudia la distribución por países, cuartiles, idioma de publicación e institución editora. Se estudia la distribución por países, cuartiles, idioma de publicación e institución editora. **Resultados:** A lo largo de los últimos diez años, el número de revistas de educación ha ido aumentando en las bases de datos JCR en todas las categorías temáticas excepto en la categoría *Education*, *Special*. Solamente 8 países tienen revistas de educación situadas en posiciones de privilegio (cuartiles 1 y 2). Estados Unidos, Reino Unido, Holanda y España,

son los cuatro países con mayor número de revistas en los cuartiles 1 y 2. En cuanto al idioma de publicación, el inglés y el español son los idiomas más utilizados. Se observa que un gran número de revistas españolas publican sus trabajos en inglés y en español. Las tres revistas españolas con mayor factor de impacto: *Comunicar*, *Revista de Educación* y *Revista de Psicodidáctica*, publican sus trabajos en inglés y español. Conclusiones: Se concluye que, en el caso de las revistas españolas de educación, la publicación de los trabajos en inglés y español, contribuye a una mayor difusión de los mismos entre la comunidad científica nacional e internacional, aumentando su visibilidad y factor de impacto.

Palabras clave: revistas de educación, factor de impacto, bibliometría, evaluación de la actividad científica, JCR.

Abstract

Introduction: The evaluation of scientific activity in Spain grants personal and social recognition to professors and researchers and conditions their academic career. Articles published in journals included in the Journal Citation Reports (JCR), are the most necessary to achieve personal accreditation and sexennial research recognition. In this paper, the situation of the education journals currently included in both JCR (SCIE and SSCI) are analysed. **Methodology:** the sample contained 340 journals included in the *Education & Educational Research, Special, Education and Psychology Educational* subject categories in the 2016 JCR Science Citation Index Expanded (SCIE) edition. The distribution by countries, quartiles, language of publication and publishing institution is studied. **Results:** Throughout the last ten years, the number of education journals has increased in the JCR databases in all of the subject categories except in *Education, Special*. Only 8 countries have education journals in privileged positions (quartiles 1 and 2). The United States of America, the United Kingdom, the Netherlands and Spain are the four countries with the highest number of journals in quartiles 1 and 2. Regarding the publication language, English and Spanish are the most used languages. It is also observed that a great number of Spanish journals publish both in English and Spanish. The three Spanish journals with greater impact factor are: *Comunicar*, *Revista de Educación* and *Revista de Psicodidáctica*, all publish in English and Spanish. **Conclusions:** Therefore it is concluded, that in the case of Spanish educational journals, the publication of the works both in English and Spanish language contributes to a greater diffusion of the papers among the national and international scientific community, increasing its visibility and impact factor.

Key words: education journals, impact factor, bibliometrics, evaluation of scientific activities, JCR

Introducción

Las revistas científicas son las principales herramientas con las que cuentan en la actualidad los investigadores para difundir los resultados de sus investigaciones. Este medio debe ser permanente, puntual, riguroso y transparente, convirtiéndose en el registro oficial y público del conocimiento. Pero, además, las publicaciones periódicas desempeñan un importante papel en el proceso de comunicación científica, ya que constituyen la principal vía para la validación del nuevo conocimiento y hacen posible su difusión dentro de la comunidad de investigadores, y desde este hacia la sociedad y sus agentes.

Las revistas científicas tienen un papel fundamental en las distintas etapas de la actividad investigadora, ya que son la vía por la que los investigadores obtienen un reconocimiento por sus contribuciones al progreso científico. La publicación de un trabajo en una revista de prestigio reconocido puede contribuir a incrementar el reconocimiento personal y social a profesores e investigadores (Carbonell y Calvó, 2009), condicionando además de forma directa la progresión en su carrera académica (Jiménez-Contreras, Torres-Salinas, Ruiz-Pérez y Delgado, 2010; Ruiz-Pérez, Martín-Martín y Delgado, 2015) y aumentar sus retribuciones económicas. En el actual contexto de globalización, la visibilidad internacional también se ha convertido en un activo fundamental para las instituciones de educación superior, ante la necesidad de estrechar lazos con otras instituciones para optimizar recursos y complementar capacidades y como estrategia de obtención de recursos y posicionamiento en el ámbito global. Ser visible internacionalmente contribuye a captar estudiantes, profesores e investigadores de prestigio y poder vincularse a otros socios para poder solicitar proyectos en convocatorias internacionales (De Filippo, Pandiella-Dominique y Sanz-Casado, 2017), la visibilidad trae aparejado reconocimiento y recursos (Docampo, 2008).

Los estudios bibliométricos sobre la producción y el impacto de las publicaciones científicas constituyen el principal medio empleado por los científicos para conocer la calidad de sus investigaciones, así como la mayor o menor relevancia de las instituciones a los que están vinculados (Cronin, 1984). La evaluación de la producción científica y de la ciencia en general ha estado tradicionalmente limitada y centrada en revistas científicas (Giménez-Toledo, Mañana-Rodríguez y Tejada-Artigas,

2015), sin embargo, los trabajos sobre el consumo de información han puesto de manifiesto la existencia de marcadas diferencias entre los hábitos de publicación y citación entre los científicos de las diferentes disciplinas (Archambault y Larivière, 2010; Osca-Lluch, Veyrat y Morales, 2013; Dorta-González y Ramirez-Sánchez, 2014). Uno de los aspectos que más diferencia a unas disciplinas de otras es el relativo al vehículo de difusión de los resultados de las investigaciones. Todos los estudios basados en los cómputos de citas coinciden en afirmar que en las ciencias experimentales la información se transmite principalmente a través de artículos de revista, que suman en ella porcentajes de citas superiores al 80 por ciento, seguidos de lejos por los libros, con tantos por ciento en torno a 10. Por el contrario, en las ciencias sociales y las humanidades y, por distintos motivos, también en las aplicaciones prácticas, predominan los libros, que reúnen entre el 50 y el 65 por ciento de las citas, en tanto que los artículos de revista tienen un peso relativamente modesto, que se manifiesta en porcentajes comprendidos entre el 10 y el 35 (Broadus, 1971). La relevancia de los libros es indiscutible y la ausencia de bases de datos internacionales exhaustivas que cubran los elementos y la información necesaria para la evaluación de este tipo de publicaciones ha instado a varios países europeos a desarrollar sistemas de información personalizados para el registro de libros académicos (Giménez-Toledo et al., 2016; Giménez-toledo, Mañana-Rodríguez y Sivertsen, 2017). Este distinto comportamiento a la hora de publicar tiene, por lo general, su correspondiente reflejo en las pautas de citación de los autores. De los trabajos realizados sobre los humanistas, se han obtenido unas características comunes a estos investigadores entre las que cabe mencionar que se observa que utilizan preferentemente los canales formales para actualizar su conocimiento, y dentro de ellos, suelen consultar más monografías que publicaciones periódicas. Los canales informales tienen una gran importancia, aunque los emplean fundamentalmente para mantener contactos personales con colegas, aunque tienen tendencia a trabajar de forma aislada, por lo que existen relativamente pocos grupos de colaboración en esta área y en general, tienen una gran capacidad idiomática, por lo que pueden hacer uso de documentos escritos en diferentes idiomas (Osca-Lluch, Veyrat y Morales, 2013). Sin embargo, la evaluación de la producción científica y de la ciencia en general ha estado tradicionalmente limitada y centrada en revistas científicas (Giménez-Toledo, Mañana-Rodríguez y Tejada-

Artigas, 2015), y a diferencia de lo que ocurre con las revistas, son pocos los instrumentos e indicadores que existen para evaluar la calidad de las monografías (Giménez-Toledo y Tejada-Artigas, 2015).

En los últimos años ha habido grandes cambios en cuanto a edición, valoración y jerarquización del mérito de las revistas (Santos y Fernández-Ríos, 2016). La citación de artículos es el indicador de visibilidad más importante a la hora de medir el impacto académico de las revistas y del conocimiento allí publicado. Los indicadores de impacto son valores que se calculan a partir de la cantidad de citaciones registradas por los trabajos publicados en una revista, los cuales son generalmente aceptados como medidas que indican la importancia, el impacto, el prestigio o la visibilidad de una revista con respecto a las otras de su mismo campo científico y que sirven para establecer rankings de las revistas indexadas en las propias bases de datos, de acuerdo con indicadores que ellos mismos calculan.

El indicador más reconocido y frecuentemente utilizado para evaluar la calidad de las publicaciones científicas y medir el impacto de una revista es el factor de impacto de los Journal Citation Reports (JCR), un indicador bibliométrico actualmente de la empresa Clarivate Analytics, que refleja el promedio de citación de artículos publicados en una revista científica indexada en el sistema Web of Science en los índices *Science Citation Index Expanded (SCIE)* o en el *Social Sciences Citation Index (SSCI)*, no en el *Arts & Humanities Citation Index (A&HCI)*. El uso del factor de impacto ha sido ampliamente criticado, como por ser una limitación para publicaciones no inglesas (González-Alcaide, Valderrama-Zurián y Aleixandre-Benavent, 2012) o porque el número de citas no mide la calidad del artículo (Velasco, Eiros, Pinilla y San Román, 2012). Con el paso del tiempo a los indicadores de impacto se les han encontrado algunos inconvenientes, lo que ha llevado a la formulación de otros nuevos que intentan siempre ser una medida confiable de la calidad, prestigio o la importancia de las publicaciones (Bormann y Daniel, 2007; Buela-Casal, 2002, 2003; Dodson, De Souza y Dos Santos, 2012; Egghe, 2006; Hirsch, 2010; Hirsch y Buela-Casal, 2014; Ramirez Martinez, Martinez Ruiz y Castellanos, 2012), no obstante, los nuevos indicadores no han tenido suerte o tanta aceptación como el de factor de impacto (FI) y tampoco están exento de críticas. Uno de los aspectos más criticados del uso del factor de impacto del JCR es la escasa presencia de publicaciones de algunos países de habla no inglesa, pese a que los

países iberoamericanos más desarrollados han llevado a cabo rigurosas políticas para potenciar las publicaciones nacionales de mayor calidad (García-Pereira y Quevedo-Blasco, 2015; Quevedo-Blasco y López-López, 2011; Rodríguez Yunta, 2010).

La implantación en España de un sistema de evaluación de la actividad investigadora ha sido muy controvertida desde sus orígenes, sobre todo desde que diera lugar a incentivos económicos mediante los conocidos tramos o sexenios de investigación (Galán, 2014). El sistema otorga reconocimiento personal y social a profesores e investigadores, condicionando además de forma directa la progresión en su carrera académica. Los profesores que no logran que se les reconozcan sus investigaciones se ven obligados a emplear la mayor parte de su jornada en dar clases, con lo que no les queda tiempo para investigar (Ruíz-Pérez, Martín-Martín y Delgado, 2015).

Los criterios recientemente publicados que indican los méritos necesarios para conseguir la evaluación de la actividad investigadora vuelven a reiterar que se valorarán preferentemente “ los artículos publicados en revistas de reconocida valía, aceptándose como tales las incluidas en los listados por ámbitos científicos del Journal Citation Reports (Social Sciences Edition) y Journal Citation Reports (Science Edition) de la «Web of Science” y entre otras consideraciones, también se indica que “se valorará desfavorablemente la reiterada publicación de trabajos en revistas o editoriales pertenecientes o asociadas al mismo organismo donde el solicitante realiza su investigación, así como que las distintas aportaciones se dupliquen o resulten iterativas sin efectiva innovación. También se valorará desfavorablemente el que más de una contribución forme parte de un mismo libro o número de revista. Igualmente, cuando el trabajo muestre un nivel alto de autocitas”. (CNEAI, 2017).

Por todas estas razones, conocer las revistas de una determinada área científica o disciplina, resulta fundamental para los investigadores, con el fin de que puedan seleccionar qué revistas son las más indicadas para publicar sus trabajos y, en los últimos años, se han realizado diferentes estudios cuyo objetivo es el análisis de la difusión y posición de las revistas científicas en diferentes bases de datos (Castillo y Carretón, 2010; García-Pereira y Quevedo-Blasco, 2015; González-Sala y Osca-Lluch, 2016; Olivas-Avila, Musi-Lechuga, Quevedo-Blasco y Luna-Hernández, 2012; Quevedo-Blasco, 2013; Quevedo-Blasco y López-López, 2011; Repiso, Jiménez-Contreras y Aguaded, 2017).

Si las revistas, como medios de publicación y difusión de la investigación, tienen una gran relevancia, no lo son menos las plataformas o bases de datos donde estas son indexadas. El objetivo general de este trabajo es caracterizar las revistas de educación recogidas en los Journal Citation Reports (JCR). El análisis se centra en dos aspectos fundamentales, la posición de las revistas y distribución según los diferentes países, y los aspectos editoriales con el propósito de conocer las revistas de excelencia en esta disciplina científica.

Método

Unidades de análisis y materiales

Se analizaron todas las revistas indexadas en las categorías temáticas *Education & Educational Research*, *Special Education* y *Psychology Educational* de la edición del Journal Citation Reports (JCR) del Social Sciences Citation Index (SSCI) y de la categoría temática *Education, Scientific Disciplines* de la edición del Journal Citation Reports (JCR) del Science Citation Index Expanded (SCIE) del 2016 (correspondiente al año 2017).¹

En las bases de datos JCR (SCIE y SSCI), una misma revista puede estar incluida en una o en las dos bases de datos JCR y además, también puede estar indexada en una o varias categorías temáticas en cada una de las bases de datos. De las cuatro categorías estudiadas: *Education & Educational Research*, es la que tiene un mayor número de revistas relacionadas con toda la gama de trabajos sobre educación, desde teóricos hasta aplicados. En ella se indexan recursos sobre pedagogía y metodología, así como sobre la historia de la educación, la lectura, los estudios curriculares, la política educativa y la sociología y la economía en educación, así como sobre el uso de ordenadores en el aula. La categoría *Education, Special* incluye revistas que publican trabajos relacionados con la educación y el desarrollo de personas con necesidades especiales, incluidos los superdotados y aquellos con discapacidades de aprendizaje. *Psychology, Educational* incluye publicaciones sobre psicología educativa,

⁽¹⁾ Los datos para este estudio fueron recogidos en el mes de enero del 2018.

comportamiento creativo, ciencias de la educación, investigación de la lectura y psicología escolar. Por último, la categoría *Education, Scientific Disciplines*, incluye revistas que cubren todos los recursos educativos en las diferentes disciplinas científicas.

Se han revisado todas las revistas incluidas en las diferentes categorías relacionadas con la educación, con el fin de detectar el solapamiento existente entre las mismas. Una vez eliminadas las revistas duplicadas, es decir, las que han sido clasificadas en más de una categoría temática, se han identificado un total de 340 revistas.

Procedimiento

De cada una de las revistas se han analizado la distribución por categorías temáticas, países, tipología institucional, idioma y cuartiles que ocupan en la base de datos JCR².

Resultados

Las revistas relacionadas con la disciplina de Educación, indexadas en la base de datos JCR-2016 pueden estar clasificadas temáticamente, en una o más de las cuatro categorías mencionadas anteriormente. La categoría *Education & Educational Research*, es la que incluye en la actualidad un mayor número de revistas (236). El segundo lugar, en cuanto a número de revistas recogidas, lo tiene la categoría *Psychology Educational*, con 58 revistas. El tercer y cuarto puesto lo ocupan las categorías *Education, Scientific Disciplines* y *Education, Special*, con 41 y 38 revistas respectivamente.

Hay que destacar que, a lo largo de los últimos años, el número de revistas de educación ha ido aumentando en las bases de datos JCR en todas las categorías temáticas excepto en la categoría *Education, Special*, que baja ligeramente en la edición del último JCR, el del año 2016, correspondiente a la edición del año 2017, tal como se muestra en la

⁽²⁾ Toda la información relativa a las revistas incluidas en las bases de datos JCR (SCIE y SSCI) se puede descargar de ISI Web of Knowledge en formato PDF, CVS o XLS, en <https://jcr.incites.thomsonreuters.com/>

tabla I, y, por el contrario, la categoría que presenta un mayor crecimiento es la de *Education & Educational Research*. Cuando se observa que el factor de impacto medio de las diferentes categorías temáticas a lo largo de todo el periodo analizado, se observa que los factores más altos no siempre coinciden con las categorías que tienen un mayor número de revistas. También llama la atención que dentro de una misma categoría tampoco hay una equivalencia entre mayor número de revistas indexadas y mayor factor de impacto de esa categoría. En las categorías *Education, Scientific Disciplines, Education, Special* y *Psychology, Educational* es en los años 2013 y 2014 cuando se obtienen unos valores de factor de impacto más elevado en cada una de estas categorías, mientras que en el caso de *Education & Educational Research*, son el 2014 y 2016.

TABLA I. Distribución del número de revistas en JCR por años, categoría temática y factor de impacto (FI) medio de la categoría temática

Años	Education & Educational Research	FI	Education, Scientific Disciplines	FI	Education, Special	FI	Psychology Educational	FI
2007	105	0,548	23	0,663	29	0,529	38	0,813
2008	113	0,711	24	0,883	30	1,000	42	1,037
2009	139	0,723	27	0,779	30	1,117	44	0,982
2010	184	0,906	33	1,220	36	1,373	50	1,416
2011	206	0,708	33	0,902	37	0,906	51	1,116
2012	219	0,644	34	0,945	36	0,914	51	1,100
2013	219	0,914	36	1,430	37	1,694	53	1,580
2014	224	0,922	37	1,352	39	1,447	55	1,687
2015	231	0,896	40	1,211	39	1,000	57	1,219
2016	236	1,107	41	1,329	38	1,168	58	1,406

Una misma revista puede estar clasificada en diferentes categorías temáticas y, además, ocupar diferente posición y cuartil en cada una de las categorías en las que ha sido incluida. Cuando se analiza el número de revistas de educación, eliminando los solapamientos existentes entre

las cuatro categorías analizadas indexadas en el JCR 2016, se observa que el número de publicaciones es de 340.

El análisis de la distribución por países muestra que las revistas de educación incluidas en el JCR están editadas por un total de 22 países. Estados Unidos y Reino Unido son los dos países que cuentan con el 78,66% de las revistas de educación incluidas en el JCR del 2016. Solamente hay 6 países que tienen más de 5 revistas de educación indexadas actualmente en el JCR, que son Estados Unidos (156 revistas), Reino Unido (111 revistas), Holanda (22 revistas), España (9 revistas), Alemania (8 revistas) y Australia (8 revistas).

El análisis de la distribución de las revistas por cuartiles (ver tabla II) pone de manifiesto que solamente seis países son los que tienen alguna revista situada en el cuartil de mayor impacto o primer cuartil (Canadá, España, Estados Unidos, Holanda, Nueva Zelanda y Reino Unido). El número de países con revistas incluidas en el cuartil 2, es de siete (Australia, España, Estados Unidos, Holanda, Nueva Zelanda, Portugal y Reino Unido). Por lo tanto, podemos ver que de los 22 países con revistas de educación indexadas en JCR, solamente hay 8 países con revistas que están incluidas en posiciones de privilegio (cuartiles 1 y 2). Como se puede observar, el mayor número de países tiene sus revistas posicionadas en el cuartil 4, excepto Bélgica, Canadá, Dinamarca, México y Portugal.

TABLA II. Distribución de las revistas por países y cuartiles*

Países	Q1	Q2	Q3	Q4	Total revistas
Alemania			3	5	8
Australia		1	3	4	8
Bélgica			1		1
Brasil				1	1
Canadá	1		1		2
Corea del Sur				2	2
Croacia				1	1
Dinamarca			1		1
España	1	2	2	4	9

Estados Unidos	53	42	39	22	157
Filipinas				1	1
Holanda	6	9	4	3	18
India				1	1
Italia				1	1
Lituania				1	1
México			1		1
Nueva Zelanda	1	1	1	1	4
Portugal		1			1
Reino Unido	35	20	31	25	111
Rusia				1	1
Sudáfrica			2	2	4
Turquía			1	2	3
Total	97	76	90	77	340

* Datos del JCR 2016. Solo se muestra una revista por cuartil. Cuando una misma revista está incluida en más de una categoría temática, se ha contabilizado en la tabla solamente el cuartil en la que está mejor posicionada.

Si centramos la mirada en las revistas españolas de educación incluidas en las bases de datos JCR (SSCI y SCIE), se observa que las primeras revistas científicas españolas de esta disciplina no aparecen hasta el siglo XX, concretamente en la década de los 40. Hay que esperar hasta la década de los 80 para encontrar un aumento relevante en las escasas publicaciones periódicas focalizadas en educación. Como precursoras destacan dos revistas: la *Revista Nacional de Educación* (1941), que posteriormente adoptó el título de *Revista de Educación* (1952). Desde su origen es editada por el Ministerio de Educación y en la actualidad es una revista reconocida en toda la comunidad académica y científica (Ruiz-Corbella, Galán y Diestro, 2014). Desde el año 2010 se encuentra indizada en la base JCR del SSCI, en el área temática *Education & Educational Research*. La segunda es la *Revista Española de Pedagogía* (1943), promovida por el Instituto San José de Calasanz de Pedagogía del Consejo Superior de Investigaciones Científicas (CSIC), y que en la actualidad es editada por la Universidad Internacional de la Rioja (Ruiz-Corbella, Galán y Diestro, 2014). Esta publicación se encuentra también

incluida en el área temática *Education & Educational Research* de la base de datos JCR (SSCI) desde el año 2007.

Las otras revistas españolas de esta disciplina que están incluidas actualmente en el JCR son:

- *Infancia y Aprendizaje* (1978), editada por la Fundación Infancia y Aprendizaje y Taylor & Francis, que se encuentra indexada en las categorías *Psychology, Educational* y *Psychology, Developmental* del JCR (SSCI) desde el año 2008.
- *Enseñanza de las Ciencias* (1983), de la Universitat Autònoma de Barcelona, indexada en la categoría *Education & Educational Research* del JCR (SSCI) desde el año 2010.
- *Comunicar* (1994), editada por el Grupo Comunicar, que se encuentra indexada desde el año 2009 en las categorías temáticas *Education & Educational Research* y *Communication* del JCR (SSCI).
- *Cultura y Educación* (1996), editada por la Fundación Infancia y Aprendizaje y Taylor & Francis. Se encuentra incluida desde el año 2010 en la categoría *Education & Educational Research* del JCR (SSCI).
- *Revista de Psicodidáctica* (1996), editada por la Universidad del País Vasco. Indexada en la categoría *Psychology, Educational* del JCR (SSCI) desde el año 2009.
- *Educación XXI* (1998), editada por la Universidad Nacional de Educación a Distancia (UNED), indexada en la categoría temática *Education & Educational Research* del JCR (SSCI) desde el año 2011.
- *Porta Linguarum* (2004), editada por la Universidad de Granada. Se encuentra indexada en las categorías temáticas *Education & Educational Research* y *Linguistics* del JCR (SSCI) desde el año 2010.

Cuando se analiza el idioma de las publicaciones, se observa que los trabajos publicados en las revistas de educación se pueden publicar en 6 idiomas diferentes. El idioma más utilizado por las publicaciones es el inglés (94,41%, 321 revistas), el segundo lugar lo ocupan las revistas que admiten trabajos escritos en más de un idioma (multilingüaje), con un 3,82% (13 revistas). En el caso de las publicaciones que editan sus trabajos en más de un idioma, los casos más frecuentes son las que

aceptan los trabajos en inglés y español o inglés y alemán. Los otros 5 idiomas utilizados por las publicaciones son el español (0,59%, 2 revistas), el croata, el italiano, el ruso y el turco, con un porcentaje del 0,29 (1 revista) por ciento cada uno de ellos.

Cuatro grupos de instituciones publican las revistas de educación indexadas en el JCR: Las editoriales comerciales, que publican el 77,94% de las revistas, las Fundaciones y Asociaciones (11,47%), las Universidades (10,29%) y los Organismos Gubernamentales (0,29%). Dentro del grupo de editoriales comerciales, destacan Routledge Journals, Taylor & Francis LTD (98 revistas), Wiley-Blackwell (75 revistas), Sage Publications LTD (42 revistas), Springer (31 revistas) y Elsevier (21 revistas).

Con respecto al factor de impacto de las revistas, se puede apreciar que las revistas incluidas en las categorías que indexan un mayor número de revistas, *Education & Educational Research* y *Psychology Educational*, son las que presentan un mayor factor de impacto y, por el contrario, las revistas que tienen un menor factor de impacto, son las incluidas en las categorías que tienen un menor número de revistas, como son *Education Special* y *Education Scientific Disciplines*.

La revista que tienen un mayor factor de impacto es *Education Psychologist* (FI 6,257), que ocupa el primer puesto en las categorías *Education & Educational Research* y *Psychology Educational*. La segunda revista con un mayor factor de impacto es *Review of Educational Research* (FI 5,263), que esta indexada en la categoría *Education & Educational Research*. En la categoría *Education Scientific Disciplines*, la revista con un mayor factor de impacto es *Academic Medicine* (FI 5,255) y en la categoría *Education Special*, la revista con un mayor factor de impacto es *Journal of Influenza Disorders* (FI 2,714).

Discusión

Este estudio pretende ofrecer un panorama de la situación de las revistas de educación en el JCR del 2016, donde hay indexadas un total de 340 revistas clasificadas en el Journal Citation Reports (JCR) del 2016 (editado en el 2017). Las revistas juegan un importante papel en la evaluación de la actividad científica, tanto de los autores como de sus instituciones y, por ese motivo, es fundamental, que los investigadores conozcan que revistas son las más idóneas para la publicación de sus

trabajos. Los cambios legislativos introducidos desde la década de los años 80 en España dejaron patente que la evaluación y el reconocimiento del profesorado universitario pasaba necesariamente por las evidencias de sus publicaciones especialmente en forma de artículos de calidad contrastada o, dicho de otra forma, revistas con impacto. Resulta innegable que esta política evaluativa ha impulsado a las revistas españolas a adecuarse a los patrones de difusión internacional de la productividad científica para mejorar su impacto (Ruiz-Corbella, Galán y Diestro, 2014). Sin embargo, para algunos autores, estas políticas pueden provocar una auténtica perversión del sistema, ya que puede llevar a los investigadores a centrarse en diferentes estrategias para publicar y ser citado, en detrimento de otros criterios de calidad y ética científica (Buela-Casal, 2014).

Uno de los indicadores que por su relevancia internacional se ha convertido en una referencia para valorar la calidad de las publicaciones periódicas son los Journal Citation Report (JCR). En el caso de las revistas de educación incluidas en JCR, como era de esperar, dada las características de estas bases de datos, los países que tienen un mayor número de revistas son Estados Unidos y Reino Unido, sin embargo, hay que destacar la presencia de revistas pertenecientes a países como Holanda y España, que ocupan el tercer y cuarto puesto en número de publicaciones de esta disciplina.

La visibilidad de las publicaciones científicas españolas ha aumentado considerablemente y se observa que, a partir del año 2007, se aprecia un aumento significativo del número de revistas indexadas (Moreno-Pulido, López-González, Rubio-Garay, Saúl y Sánchez-Elvira-Paniagua, 2013). El análisis de las revistas de educación en JCR revela que son 9 las revistas españolas de educación incluidas en JCR y que solamente son 3 las que ocupan actualmente posiciones de privilegio (cuartiles 1 y 2).

En cuanto al idioma de publicación, el inglés se acepta en 329 de las 340 revistas, y el español es el segundo idioma de publicación más utilizado, aunque un gran número de revistas españolas publican sus trabajos en inglés y en español (7 revistas). Las tres revistas españolas de mayor factor de impacto, como son *Comunicar* (cuartil 1 en las categorías *Education & Educational Research* y *Communication*), *Revista de Educación* (cuartil 2 en la categoría *Education & Educational Research*) y *Revista de Psicodidáctica* (cuartil 2 en la categoría *Psychology Educational*)

publican sus trabajos en inglés y español, lo que contribuye a dar una mayor difusión entre la comunidad científica nacional e internacional.

El estudio del factor de impacto de las revistas revela que en las categorías donde hay un mayor número de revistas indexadas, no siempre son las que obtienen un factor de impacto más alto, sino que está relacionado con las citaciones que reciben los trabajos. Algunos estudios revelan que cada vez se cita más, pero que aparecen algunos elementos que llaman la atención como crecimientos en la tasa de autocitación que seguramente tengan más que ver con estrategias editoriales en pos de un mayor índice de impacto que un aumento real de la influencia de estas revistas (Fernández-Quijada, 2010).

Como conclusión de este trabajo se puede afirmar que la implantación de una política evaluativa sustentada en indicadores internacionales ha contribuido a mejorar la calidad y visibilidad de las revistas españolas de educación. España ocupa el cuarto lugar, en el ranking de países que tienen revistas de educación situadas en los cuartiles 1 y 2 en el JCR, tras Estados Unidos, Reino Unido y Holanda, que son los tres países que ocupan los tres primeros puestos. De las 9 revistas españolas indexadas actualmente en JCR 3 de ellas, están situadas ocupando posiciones de privilegio (cuartiles 1 y 2), lo que viene a revelar la aceptación que tienen los trabajos publicados en las mismas entre la comunidad científica y una de las conclusiones de este trabajo es que es recomendable, la publicación de los trabajos en los dos idiomas (español e inglés), ya que contribuye a mejorar la difusión y visibilidad de los trabajos publicados también entre la comunidad hispana. Se observa que las revistas españolas de educación cada vez tienen más presencia en los JCR, compartiendo los resultados de Quevedo-Blasco y López-López (2011) y García-Pereira y Quevedo-Blasco (2015), sería deseable que estas revistas españolas no solo aumentarían su factor de impacto, sino que otras nuevas fueran indexadas en estas bases de datos, con el fin de evitar fugas de trabajos a revistas multidisciplinarias o de otras áreas más generales, cuyas revistas cuentan con un mayor factor de impacto.

Referencias bibliográficas

- Archambault, E. y Larivière, V. (2010). *The limits of bibliometrics for the analysis of the social sciences and humanities literature*. En: 2010 World Social Science Report: Knowledge Divides. Paris: UNESCO, 251-254.
- Bormann, L. y Daniel, H.D. (2007). Convergent validation of peer review decisions using the h index-Extent of and reason for type I and type II errors. *Journal of Informetrics*, 1 (3), 204-213.
- Broadus, R.N. (1971): The Literature of the Social Science: A Survey of Citation Studies. *International Social Science Journal*, 23, 236-243.
- Buela-Casal, G. (2002). Evaluación de la investigación científica: “El criterio de la mayoría”: El factor de impacto, el factor de prestigio y los “diez mandamientos para incrementar las citas”. *Análisis y Modificación de Conducta*, 28, 119, 455-476.
- Buela-Casal, G. (2003). Evaluación de la calidad de los artículos científicos: propuesta del factor de impacto ponderado y de un índice de calidad. *Psicothema*, 15, 1, 23-35.
- Buela-Casal, G. (2014). Pathological publishing: a new psychological disorder with legal consequences? *The European Journal of Psychology Applied to Legal Context*, 6(2), 91-97.
- Carbonell, X. y Calvó, N. (2009). Las revistas españolas de psicología. Cómo elegir la revista donde publicar. *Anales de Psicología*, 25, 209-216.
- Castillo, A. y Carretón, M.C. (2010). Investigación en comunicación. Estudio bibliométrico de las revistas de Comunicación en España. *Comunicación y Sociedad*, 2, 289-327.
- CNEAI (2017). Resolución de 23 de noviembre de 2017, de la Comisión Nacional Evaluadora de la Actividad Investigadora, por la que se publican los criterios específicos aprobados para cada uno de los campos de evaluación. BOE núm. 292, de 1/12/2017, p. 117060-117080.
- De Filippo, D., Pandiella-Dominique, A. y Sanz-Casado, E. (2017). Indicadores para el análisis de la visibilidad internacional de las universidades españolas. *Revista de Educación*, 376, 163-199. DOI: 10.4438/1988-592X-RE-2017-376-348.
- Docampo, D. (2008). Rankings internacionales y calidad de los sistemas universitarios. *Revista de Educación*, número extraordinario, 149-176.

- Dodson, M.V., De Souza Duarte, M. y Dos Santos Dias, L.A. (2012). SP-index: The measure of the scientific production of researchers. *Biochemical and Biophysical Research Communications*, 425 (4), 701-702.
- Dorta-González, P. y Ramirez-Sánchez, M. (2014). Producción e impacto de las instituciones españolas de investigación en Arts & Humanities Citation Index (2003-2012). *Arbor*, a191. Doi: <http://dx.doi.org/10.3989/arbor.2014.770n6012>
- Egghe, L. (2006). Theory and practice of the g-index. *Scientometrics*, 69 (1), 131-152.
- Fernández-Quijada, D. (2010). El perfil de las revistas españolas de comunicación (2007-2008). *Revista Española de Documentación Científica*, 33 (4), 553-581.
- Galán, A. (2014). Evaluación de la producción científica en educación. ¿Una nueva crisis? *Bordón*, 66 (2), 7-10.
- García-Pereira, S. y Quevedo-Blasco, R. (2015). Análisis de las revistas iberoamericanas de Psicología y Educación indexadas en el Journal Citation Reports del 2013. *European Journal of Education and Psychology*, 8, 85-96.
- Giménez-Toledo, E., Mañana-Rodríguez, J. y Sivertsen, G. (2017). Scholarly book publishing: Its information sources for evaluation in the social and humanities. *Research Evaluation*, 26(2), 91-101.
- Giménez-Toledo, E., Mañana-Rodríguez, J., Engels, T.C., Ingwersen, P., Polonen, J., Sivertsen, G., Verleysen, F.T. y Zuccala, A.A. (2016). Taking scholarly books into account. Current developments in five European countries. *Scientometrics*, 107, 685-689.
- Giménez-Toledo, E., Mañana-Rodríguez, J. y Tejada-Artigas, C.M. (2015). Revisión de iniciativas nacionales e internacionales sobre evaluación de libros y editoriales. *El Profesional de la Información*, 24, 705-716.
- Giménez-Toledo, E. y Tejada-Artigas, C. M. (2015). Proceso de publicación, calidad y prestigio de las editoriales científicas en educación. *Educación XX1*, 18(1), 17-44.
- González-Sala, F. y Osca-Lluch, J. (2016). Análisis de las publicaciones españolas en la categoría Psychology Educational de la Web of Science durante el período 2004-2013. *Aula Abierta*, 44, 46-54.
- Hirsch, J.E. (2010). An index to quantify an individual's scientific research output that takes into account the effect of multiple coauthorship. *Scientometrics*, 85 (3), 741-754.

- Hirsch, J.E. y Buela-Casal, G. (2014). The meaning of the h-index. *International Journal of Clinical and Health Psychology*, 14, 2, 161-164.
- Jiménez-Contreras, E., Torres-Salinas, D., Ruiz-Pérez, R. y Delgado López-Cózar, E. (2010). Investigación de excelencia en España: ¿protagonistas o papeles secundarios? *Medicina Clínica*, 134 (2), 76-81.
- Moreno-Pulido, A., López-González, M.A., Rubio-Garay, F., Saúl, L.A. y Sanchez-Elvira.Paniagua, A. (2013). Evolución de las revistas españolas de ciencias sociales en el Journal Citation Reports (2006-2010) y su valoración como indicio de calidad en la normativa evaluadora. *Revista Española de Documentación Científica*, 36 (3), e014. doi: <http://dx.doi.org/10.3989/redc.2013.3.987>
- Olivas-Avila, J.A., Musi-Lechuga, B., Quevedo-Blasco, R. y Luna-Hernández, J.R. (2012). Índice de internacionalidad de las revistas iberoamericanas de psicología en el Journal Citation Report (2011). *Revista Latinoamericana de Psicología*, 44 (3), 175-183.
- Osca-Lluch, J., Veyrat, A. y Morales, J. (2013). El consumo de información en humanidades. *Arbor*, 189,760, a026. Doi: <http://dx.doi.org/10.3989/arbor.2013.760n2012>
- Quevedo-Blasco, R. (2013). Revistas iberoamericanas de psicología indexadas en el Journal Citation Reports de 2011. *Revista Mexicana de Psicología*, 30 (1), 1-10.
- Quevedo-Blasco, R. y López-López, W. (2011). Situación de las revistas iberoamericanas de psicología en el Journal Citation Reports de 2011. *Universitas Psychologica*, 10 (3), 955-965.
- Repiso, R., Jiménez-Contreras, E. y Aguaded, I. (2017). Revistas iberoamericanas de educación en Scielo Citation Index y Emerging Source Citation Index. *Revista Española de Documentación Científica*, 40 (4): e186. doi://dx.doi.org/10.3989/redc.2017.4.1445
- Rodriguez-Yunta, L. (2010). Las revistas iberoamericanas en Web of Science y Scopus: visibilidad internacional e indicadores de calidad. En: J. Rios Ortega, coord. *VII Seminario Hispano-Mexicano de Investigación en Bibliotecología y Documentación* (pp.347-363). México: UNAM.
- Ruiz-Corbella, M., Galán, A. y Diestro, A. (2014). Las revistas científicas de educación en España: evolución y perspectivas de futuro. *Relieve. Revista Electrónica de Investigación y Evaluación Educativa*, 20 (2). Doi: [10.7203/relieve.20.2.4361](http://dx.doi.org/10.7203/relieve.20.2.4361)

- Ruiz-Pérez, R., Martin-Martin, A. y Delgado López-Cózar, E. (2015). Las revistas universitarias en el marco de los criterios de evaluación de la actividad investigadora en España. *Revista Española de Documentación Científica*, 38 (2): e081.doi: <http://dx.doi.org/10.3989/redc.2015.2.1191>
- Ramirez Martinez, D.C., Martinez Ruiz, L.C. y Castellanos Dominguez, O.F. (2012). *Divulgación y difusión del conocimiento: las revistas científicas*. Bogotá: Universidad Nacional de Colombia.
- Santos Rego, M.A. y Fernández-Rios, L. (2016). El factor de impacto y el futuro de las revistas académicas. El riesgo de patologización. *Innovación Educativa*, 16, 35-51.
- Velasco, B., Eiros, J.M., Pinilla, J.M. y San Román, J. A. (2012). La utilización de los indicadores bibliométricos para evaluar la actividad investigadora. *Aula Abierta*, 40 (2), 75-84.

Información de contacto: Julia Haba-Osca. Universitat de València. Facultat de Filologia, Traducció i Comunicació, Departament de Filologia Anglesa i Alemanya. Avda. Blasco Ibáñez, 21, 46010, Valencia. E-mail: julia.haba@uv.es

Evaluación del Conocimiento Científico en Maestros en formación inicial: el caso de la Comunidad Valenciana

Assessment of pre-service teachers' Science knowledge: the case of Valencian Community in Spain

DOI: 10.4438/1988-592X-RE-2019-383-404

José Javier Verdugo-Perona

Florida Universitaria

Joan Josep Solaz-Portolés

Vicente Sanjosé

Universitat de València

Resumen

La alfabetización científica es en uno de los objetivos principales de la educación obligatoria, pero difícilmente se puede conseguir esta alfabetización si los profesores de enseñanza primaria tienen un pobre conocimiento científico (Appleton, 2003). Algunos estudios sugieren que, en España, el profesorado de educación primaria presenta carencias en su formación científica (Cañal, 2000 y 2008; Porlán, 2010). Por este motivo, se desarrolla un estudio sobre el conocimiento sobre ciencia básica de los maestros en formación en las universidades públicas de la Comunidad Valenciana, con suficiente validez externa (error de muestreo cercano a 5%). Se evalúa independientemente el conocimiento sobre conceptos básicos y sobre procesos elementales de la actividad científica, tanto antes de recibir formación como al finalizar la formación prevista en el Grado de Maestro de Primaria. La titulación de acceso al Grado y la especialidad de los estudios pre-grado, son consideradas para describir y analizar los resultados. Complementariamente, se estudia el posible efecto de las diferencias entre universidades en las horas de ciencias ofertadas. Los hallazgos muestran mejoras significativas, aunque de pequeño tamaño, en el conocimiento analizado. También evidencian necesidades formativas no suficientemente bien atendidas en el Grado universitario.

Palabras clave: Enseñanza de las ciencias; conceptos científicos; procesos de la ciencia; maestros/as en formación; evaluación.

Abstract

Scientific literacy is one of the main objectives in compulsory education, but this literacy can hardly be achieved if primary teachers have a low level of scientific knowledge (Appleton, 2003). In Spain, some studies suggest that primary education teachers show deficiencies in their scientific training (Cañal, 2000 y 2008, Porlán, 2010). For this reason, an assessment of pre-service elementary school teachers' knowledge on basic science is developed with enough external validity (sampling error about 5%). The scope of the study is the Valencian Community in Spain. Conceptual and procedural knowledge are assessed independently before as well as after the science subjects in the specific Degree of pre-service teacher training at University. Prior studies and their branch are taken into account in some descriptions and analyses. The possible effects of the more or less hours devoted to science by the three public universities in the considered Region, are also studied. Results show significant improvements in the levels of conceptual and procedural knowledge along the Degree, but of little effect size. Outcomes suggest that some teachers' educational needs are not yet sufficiently met at the considered universities.

Keywords: Science Education; Science Concepts; Science Processes skills; Preservice teachers, Assessment.

Introducción

La alfabetización científica se ha convertido en uno de los objetivos principales de la educación obligatoria y, de hecho, es un concepto mencionado en las últimas reformas educativas. Entre sus diferentes finalidades (Acevedo, 2004) está que los ciudadanos desarrollen un espíritu crítico sobre conocimientos científicos básicos que les permita participar en la toma de decisiones, sobre todo en aquellos asuntos socio-científicos que afectan a toda la sociedad (Gil y Vilches, 2006). El Consejo Nacional de Investigación (NRC), organismo de las academias nacionales de EEUU, define la alfabetización científica (AC en adelante) como el conocimiento y la comprensión de los conceptos y procesos científicos requeridos para la toma de decisiones personales, la participación en

asuntos cívicos y culturales y la productividad económica (NRC, 1996), y establece la indagación como el mejor proceso para alcanzar esos objetivos. Ello coincide en gran medida con Hurd (1998), para quien la AC se concibe como *“una competencia científica requerida para el pensamiento racional sobre la ciencia en relación a asuntos personales, sociales, políticos, problemas económicos y aquellas cuestiones con las que probablemente se encuentre a lo largo de la vida”* (p. 410).

Durante la educación primaria se deben sentar las bases para una adecuada formación de los ciudadanos (Lloyd, Smith, Fay, Khang, Wah y Sai, 1998), y ello hace este periodo particularmente relevante para la meta de aumentar la AC de toda la ciudadanía. Sin embargo, difícilmente se puede conseguir esta alfabetización si los profesores de enseñanza primaria tienen un pobre conocimiento científico (Appleton, 2003) y, consecuentemente, un conocimiento didáctico del contenido inapropiado para enseñar ciencias (Kang, 2007). Este aspecto se ve confirmado, por ejemplo, en diversas investigaciones que ponen de manifiesto que los profesores de primaria de ciencias sostienen ideas sobre algunos conceptos científicos que difieren de las defienden la comunidad científica (errores conceptuales) y, además, coinciden con las que (erróneamente) sostiene su propio alumnado (Krugar, Summers y Palacio, 1990; Harlen y Holroyd, 1997; Papageorgiou, Stamovlasis y Johnson, 2010; Summers y Krugar, 1992). Una deficiente formación científica de los maestros es causa de inseguridad y falta de confianza a la hora de abordar la enseñanza de las ciencias con los niños en las aulas de Primaria (Jarvis y Pell, 2004; García-Ruiz y Sánchez, 2006; Murphy y Smith, 2012) y, probablemente, genera en los niños actitudes negativas hacia la ciencia (Harlen y Holroyd, 1997; Jurievi, Glaar, Puko y Devetak, 2008).

Todo parece indicar que si se pretende mejorar la eficacia del proceso de alfabetización científica de los ciudadanos la formación del profesorado debería ser revisada, comenzando por la de los/as maestros/as de primaria. En particular, se debería pensar, en garantizar que sus niveles de dominio de conocimiento científico básico son adecuados.

En el ámbito español Cañal (2000 y 2008) y Porlán y colaboradores (2010) hacen referencia a las carencias en la formación científica de los maestros, así como a sus concepciones alternativas en diversos temas científicos. Subrayan la imposibilidad de construir, en este caso, un

conocimiento didáctico del contenido adecuado para poder conducir con éxito los procesos de enseñanza/aprendizaje en las aulas de primaria.

No obstante, los estudios realizados con maestros españoles, tanto en ejercicio como en formación, focalizan su atención en las concepciones sobre la naturaleza de la ciencias y la enseñanza/aprendizaje de las ciencias (Guisasola y Morentin, 2007; Martínez-Chico, Jiménez Liso y López-Gay, 2015; Mellado Jiménez, 1996); a propuestas formativas (Criado y García-Carmona, 2011; García Barros, Martínez Losada, Vega y Mondelo, 2000; Rivero et al., 2013; y a estudios que recogen la (auto)valoración de su experiencia formativa inicial en la Universidad (Cantón, Cañón y Ariás, 2013; García Barros, Bugallo y Fuentes, 2013). No se encuentran estudios relativamente recientes donde se evalúen sus conocimientos científicos básicos, aunque sí se pueden encontrar estudios donde se evalúan sus conocimientos matemáticos (Gómez y Gutiérrez-Gutiérrez, 2014; Sáenz Castro, 2007).

Ante la inexistencia datos actualizados sobre los niveles de conocimiento científico básico de los maestros de primaria en España, parece interesante abordar un estudio empírico que permita evaluarlo. En el presente trabajo, centraremos la atención en los maestros y maestras en formación inicial para, más adelante, abordar a los/las profesionales en ejercicio.

Objetivos de investigación

El objetivo principal de este trabajo es la evaluación de los niveles de conocimiento conceptual y procedimental en ciencia básica que presentan los y las estudiantes de la Comunidad Valenciana antes y después de su formación en el Grado de Maestro de Primaria. La intención fue que el estudio alcanzase suficiente validez externa y orientase una posible toma de decisiones sobre la suficiencia y adecuación de la formación ofrecida actualmente en ciencias básicas a los/las maestros/as de Primaria.

Asociado con este objetivo principal está la preocupación por averiguar si la formación en ciencias ofrecida en el Grado de Maestro de Primaria es suficiente y ayuda a mejorar significativamente dichos niveles, al menos de aquellas personas que los tuvieran realmente bajos al comienzo de dicho Grado. Algunos factores pueden condicionar tanto los niveles iniciales

de acceso al Grado como los finales. Entre ellos, consideraremos el tipo de estudios previos o titulación de acceso al Grado, y la especialidad de dichos estudios previos (ciencias experimentales en comparación con otras opciones de ciencias sociales y humanidades). Algunos análisis sobre conocimientos científicos al acceder a la universidad muestran resultados muy pobres y poco dependientes del título pre-universitarios de los estudiantes (Pérez de Landazábal, Benegas, Cabrera, et al., 2010). Sin embargo, estos estudios involucran sólo estudiantes de la especialidad de ciencias experimentales. En el caso presente, los estudiantes presentan heterogeneidad tanto en el tipo (título pre-grado) como en la especialidad (ciencias experimentales-otros) de los estudios anteriores al Grado. Por ejemplo, Glaser (1995; ver también Bransford, Brown y Cocking, 1999) identifican hallazgos en investigación que nos informan sobre la mejor estructura y diseño de entornos de aprendizaje. Uno de ellos sostiene: *“El conocimiento previo relevante y la intuición del aprendiz son... una fuente importante de destrezas cognitivas que puede sostener y conducir el nuevo aprendizaje...”* (p. 18. Traducido del inglés original).

También la cantidad de formación en ciencias ofrecida en el Grado podría condicionar los niveles finales alcanzados. No todas las universidades ofertan la misma cantidad de créditos obligatorios en ciencias experimentales, ya que la normativa permite a cada universidad ajustar el currículo a sus necesidades o metas. La importancia de este objetivo está en la reclamación secular de más carga lectiva para ciencias en los Grados (una reclamación que se hace en todas las áreas, por otro lado). Tener datos concretos y fiables sobre este asunto es crucial para la toma de decisiones en las universidades, especialmente cuando el aumento de créditos en una materia implica, necesariamente, la disminución en otras.

La Evaluación del conocimiento científico básico

En el presente trabajo se asume que el contenido de AC puede articularse en las tres dimensiones propuestas por Miller (1989): (1) comprensión de los procesos o métodos de la ciencia para probar nuestros modelos de la realidad; (2) vocabulario básico de términos y conceptos científicos y técnicos; y (3) comprensión del impacto de la ciencia y la tecnología en la sociedad. Las dos primeras dimensiones hacen referencia al

conocimiento sobre ciencia (comprensión y aplicación), y la tercera implica su aplicación responsable en escenarios sociales de la vida diaria.

Algunas propuestas contemplan las dimensiones anteriores en instrumentos concretos. Laugksch y Spargo (1996) desarrollaron el Test para la Alfabetización Científica Básica compuesto por 472 ítems tipo verdadero-falso que cubre 240 ideas importantes en ciencias y de actitudes hacia la ciencia. El test se basa en las recomendaciones del informe *Science for All Americans* (AAAS, 1989).

Diferentes instituciones de prestigio han propuesto y llevado a cabo estudios internacionales de largo recorrido y extensión sobre cultura científica básica de los ciudadanos adultos en general. En USA, la Fundación Nacional para la Ciencia (NSF) emplea un cuestionario conformado por 9 ítems para evaluar el conocimiento factual sobre ciencias (NSB 2014). La comprensión de los procesos científicos se evalúa mediante dos preguntas sobre probabilidad y una pregunta abierta acerca del significado del estudio científico o del trabajo experimental. Unos cuestionarios muy similares usan los Eurobarómetros (Special Eurobarometer 43, 76, 224) para el conocimiento factual. La comprensión de los procesos metodológicos de la ciencia y la tecnología se evalúa mediante tres preguntas acerca del método experimental, el uso de los grupos de control y el concepto de probabilidad. En España, la Fundación Española para la Ciencia y la Tecnología (FECYT) toma como referencia el modelo europeo para sus estudios. Tras diversos cambios, el informe de 2014 (FECYT, 2015), evalúa el conocimiento científico declarativo con ítems de tipo “verdadero-falso”, y un solo ítem para el conocimiento sobre los procesos de la ciencia (importancia del grupo control en un ensayo clínico).

Por otro lado, Brossard y Shanahan (2006) se centraron en la segunda dimensión de Miller, la comprensión de los términos científicos y técnicos. Validaron 31 ítems del tipo “rellenar el hueco para completar la frase con los términos apropiados”, seleccionados entre los que aparecían usualmente en los Mass Media. Los resultados obtenidos se compararon con los obtenidos por la National Science Foundation para AC. La correlación fue aceptable cuando se controló para edad, género y nivel de educación científica previo. En la misma línea Rundgren, Rundgren, Tseng, Lin, y Chang (2010) diseñaron un instrumento (llamado SLiM) para evaluar AC, formado por 50 ítems de opción múltiple. Se incluyen

los conceptos científicos que más presencia tienen en libros de texto y periódicos en Taiwán. El instrumento mostró fiabilidad en el contexto taiwanés con estudiantes desde 7º a 10º grado.

Todos estos estudios y los instrumentos usados en ellos han inspirado el presente trabajo en el contexto español pero focalizado en futuros maestros de primaria, en lugar de hacerlo sobre la población adulta en general. Hasta donde llega nuestro conocimiento esto no se había contemplado antes en un estudio de amplia validez.

Metodo

Diseño del estudio

Para abordar los objetivos el conocimiento científico fue articulado en dos componentes: a) conocimiento del contenido, de naturaleza conceptual, y; b) conocimiento procedimental, asociado a los procesos intelectuales implicados en la ciencia. Se definieron las siguientes variables dependientes:

- (a) Conocimiento conceptual sobre ciencias a nivel de educación primaria.
- (b) Conocimiento sobre procesos científicos básicos.

En lo posible se procuró que cada componente no interfiriera en la evaluación del otro (no fuera requisito o condición para el otro). Para ello, se eligieron con cuidado los instrumentos para su evaluación (ver más adelante).

Se siguió un diseño *ex-post facto* factorial mixto, de carácter sincrónico, con un factor entre-sujetos, denominado “Fase” o Momento de la formación en ciencias durante el Grado, con dos niveles: inmediatamente antes/ inmediatamente después de toda la formación obligatoria en ciencias o en didáctica de las ciencias, dentro del Grado.

De un modo complementario, se analizaron los resultados asociados con diferentes “Bloques conceptuales” o “Categorías procedimentales” específicas del conocimiento a estudiar (factores intra-sujetos). A partir de los documentos oficiales analizados, en el conocimiento conceptual se

diferenciaron 4 bloques de contenidos: I) el entorno y su conservación¹; II) diversidad de los seres vivos; III) salud y desarrollo personal; y IV) materia y energía. En el conocimiento procedimental se diferenciaron 5 categorías de procesos básicos: I) identificar y controlar variables; II) establecer hipótesis; III) definir operacionalmente; IV) interpretar gráficas y datos; y V) diseñar experimentos.

Para responder las preguntas de investigación secundarias se consideró otros factores entre-sujetos: “Titulación de Acceso al Grado” (Bachillerato y Grado o Licenciatura universitaria vs. ciclo formativo Superior de Formación Profesional); rama o “Especialidad de los estudios anteriores al Grado”, diferenciando únicamente Ciencias experimentales del resto; “Carga lectiva” de ciencias y didáctica de las ciencias obligatoria en el Grado.

En relación con la carga lectiva obligatoria en ciencias experimentales, las tres universidades públicas que imparten el Grado en Maestro/a de Educación Primaria en la Comunidad Valenciana, son: la Universidad Jaime I (UJI de Castellón), la Universidad de Valencia (Estudios Generales) y la Universidad de Alicante. La Universidad Jaime I oferta un total de 120h presenciales en dos asignaturas obligatorias anuales: “Didáctica de la Física y la Química” (6 créditos) y “Didáctica de las Ciencias Naturales” (6 créditos), impartidas ambas en 2º curso. La misma cantidad de créditos se ofertan en la Universidad de Alicante en dos asignaturas obligatorias de 6 ECTS cada una: “Enseñanza y Aprendizaje de las Ciencias Experimentales I”, (primer cuatrimestre de 2º curso), y “Enseñanza y aprendizaje de las Ciencias Experimentales II”, (segundo cuatrimestre de 3er curso). La Universidad de Valencia (Estudio General) dedica 180 horas presenciales a la formación obligatoria en Ciencias Experimentales. Los créditos se reparten en las asignaturas “Ciencias Naturales para Maestros” (anual en 2º curso, con 9 créditos), “Didáctica de las Ciencias I: materia, energía y máquinas” (cuatrimestral en 3er curso y 4,5 créditos), y “Didáctica de las Ciencias II: medioambiente, biodiversidad y salud” (cuatrimestral en 4º curso y con 4,5 ECTS). En este caso, la primera materia se dedica a la formación científica básica que necesita un maestro de primaria, y las

⁽¹⁾ Los contenidos del bloque “el entorno y su conservación” que antes estaba en Conocimiento del Medio, está ahora dentro del bloque “el mundo en que vivimos” dentro de Ciencias Sociales, según el Decreto 108/2014 (DOGV 7311/07.07.2014) para el currículo de Educación Primaria en la Comunidad Valenciana. Dado que se trata de un contenido tanto experimental como social, hemos decidido seguir incluyéndolo en este estudio.

otras dos se centran en la formación didáctica. Esta diferenciación entre formación básica en ciencias y formación didáctica no aparece en las otras dos universidades.

Muestra

Las poblaciones consideradas para este estudio están formadas por todos los estudiantes que cursan el Grado de Maestro/a en Educación Primaria en universidades públicas de la Comunidad Valenciana y que se encuentran en los cursos donde se inicia o acaba la formación obligatoria en ciencias del Grado. El tamaño de las poblaciones se calculó a partir de los datos proporcionados por las universidades en sus páginas web oficiales². Un total de 1210 alumnos/as conforma la población en la Fase 1 (antes de la formación). De acuerdo con las propias universidades, la tasa de abandono global se estima en torno al 10%. Por tanto, la población en la Fase 2 (tras la formación), es de 1100 estudiantes aproximadamente.

El muestreo se realizó en las 3 facultades de las universidades públicas de la Comunidad Valenciana consideradas.

Las muestras asociadas a cada una de las medidas realizadas (conocimiento conceptual y conocimiento procedimental), así como el error muestral en cada caso al nivel de confianza del 95%, se presentan en la Tabla I.

TABLA I. Población, muestra y error muestral en cada Fase

Fase de la Formación	Población (N)	Conceptos		Procedimientos	
		Muestra	Error muestral	Muestra	Error muestral
1 (antes)	1210	289	5,0%	308	5,2%
2 (después)	1100	270	4,8%	286	5,0%

Fuente: elaboración propia

⁽²⁾ Universidad de Alicante: <http://cvnet.cpd.ua.es/webcvnet/planestudio/planestudiond.aspx?plan=C254#tramites>
Universidad Jaime I : <http://ujiapps.uji.es/estudis/oferta/base/graus/2015/mestre-primaria/accedir/oferta-procediment-admissio/>
Universidad de Valencia: <http://www.uv.es/uvweb/universitat/ca/estudis-grau/grau-1285846094474/Titulacio.html?id=1285847460730>

La unidad de muestreo fue el grupo-clase. En cada curso se eligió el grupo de informantes y se obtuvo información voluntaria de cada estudiante presente. Se usaron sesiones ordinarias de clase de asignaturas del área de didáctica de las ciencias experimentales. A pesar de que no se siguió un procedimiento estrictamente aleatorio para elegir los grupos de informantes, el azar se asoció con el día/hora en que se acudió a la facultad a obtener los datos, dado que diferentes grupos tienen docencia de las materias de ciencias en días/horas distintas.

Los porcentajes de alumnos presentes en los grupos de aula en cada universidad donde se tomaron los datos (alumnos presentes/alumnos matriculados en el grupo) fueron: 88% (Alicante), 91% (Castellón) y 97% (Valencia). La tasa de muestreo de grupos de estudiantes se acercó al 50% tanto antes de la formación en ciencias como después de la misma. Sin embargo, por razones de tiempo cedido por los profesores, y de cansancio de los participantes (cada instrumento consta de 30 ítems de 4 opciones), sólo se obtuvo datos bien sobre conceptos, o bien sobre procedimientos de cada estudiante. En cada grupo los cuestionarios de conceptos y de procedimientos fueron repartidos al azar entre los estudiantes.

TABLA II. Distribución de las muestras

	Fase I (antes de formación)				Fase 2 (después de formación)			
	Especialidad estudios previos				Especialidad estudios previos			
Titulación Acceso	Ciencias Exper.		Otras		Ciencias Exper.		Otras	
Bachiller o Universidad	H:10	M:33	H:57	M:147	H:12	M:36	H:36	M:137
Tec. Sup. FP o Acc. Mayores	H: 0	M:2	H:7	M: 27	H:1	M:4	H:10	M: 30

Fuente: Elaboración propia

H: hombres; M: mujeres. Casos excluidos (mortalidad experimental) N=10

Instrumentos

La medida del conocimiento de contenido conceptual se basó en un instrumento desarrollado y validado por Verdugo-Perona, Solaz-Portolés y Sanjosé (2016) para maestros españoles de Educación Primaria en formación. Un proceso de reducción de longitud, mejora y adecuación condujo a la eliminación de ítems con promedios extremos (mayor de 0,85 o menor de 0,15 sobre 1) o con baja discriminación (índice de discriminación menor de 0,2). El resultado fue un cuestionario de 30 ítems cuya fiabilidad se consideró suficiente (Fórmula nº 20 de Kurder-Richardson; $KR-20=,7$) dada la gran variabilidad en su contenido.

La característica importante de este instrumento es que permite evaluar todas las áreas científicas del currículum de Educación Primaria en España y ninguno de sus ítems requiere otro conocimiento que no sea puramente conceptual. Ello implica la memoria a largo plazo, pero se minimiza su impacto al ofrecer 4 alternativas en cada pregunta (sólo una correcta). En cada ítem solamente se implica un concepto.

Se asignó 0/1 punto por cada respuesta correcta/incorrecta. La valoración final fue la suma de los puntos en los 30 ítems (no se restó por respuesta incorrecta).

Los ítems se agrupan en cuatro bloques de contenidos, que fueron definidos a partir del análisis teórico, de documentos normativos oficiales, y de la articulación temática de los textos-fuente de donde se extrajeron los ítems (ver sección Diseño del Estudio). Las correlaciones de Pearson entre todos los pares de bloques fueron significativas ($,25 < r < ,34$; $p < ,001$).

El conocimiento procedimental de los y las estudiantes se analizó mediante un cuestionario de fácil administración ya validado (Monde-Mónica, 2005), formado por 30 ítems de elección múltiple que no favorecen ninguna disciplina específica en el estudio de las ciencias. Los ítems no requieren conocimiento conceptual para contestarlos. La puntuación global en el instrumento también se obtiene por la suma de las puntuaciones en cada uno de los ítems, asignando 0/1 a las respuestas correctas/ incorrectas. Los ítems se agrupan en 5 categorías (ver Diseño del Estudio). Las distintas categorías, a excepción de “diseñar experimentos”, correlacionaron también de forma significativa entre sí ($,28 < r < ,30$; $p < ,001$ para todos los pares). “Diseñar experimentos” sólo apareció algo relacionado con “definir operacionalmente” ($r = ,10$; $p = ,028$).

El instrumento en su conjunto alcanzó también fiabilidad suficiente en nuestra muestra ($KR-20=,7$).

En ambos instrumentos, y debido probablemente al escaso número de ítems implicados, ni en la puntuación de cada bloque conceptual por separado (entre 6 y 9 ítems), ni en la de cada categoría procedimental por separado (entre 3 y 8 ítems) se alcanzó consistencia interna suficiente ($KR-20 < ,6$). Por tanto, las puntuaciones de cada bloque y las de cada categoría procedimental son sólo orientadoras del dominio de los participantes en cada temática o procedimiento particular. Los resultados y conclusiones que se deriven de dichas puntuaciones parciales tendrán fiabilidad limitada, todavía no suficientemente contrastadas y necesitadas de posteriores replicaciones. Sin embargo, los resultados y análisis asociados con puntuaciones globales de los instrumentos (colapsando bloques de contenido o categorías de procedimientos) sí tendrán fiabilidad suficiente.

Cuando se requirió, las puntuaciones de los bloques conceptuales o categorías procedimentales se normalizaron al rango 0-1 para poder ser comparadas entre sí, con independencia del número de ítems implicados.

El Anexo muestra ejemplos de ítems de ambos instrumentos.

Procedimiento

Tras obtener los permisos oportunos en cada una de las Facultades implicadas, el primer autor acudió personalmente a cada grupo intacto para solicitar la participación voluntaria y, en su caso, obtener los datos de alumnos en situación inmediatamente antes de comenzar su formación en ciencias y tras finalizarla. En cada universidad, la toma de datos se realizó el mismo día para todos los participantes. No se tomaron datos en otros momentos intermedios de la formación en ciencias.

La toma de datos se realizó durante el horario de clases ordinarias situadas en la primera o segunda sesión de cada turno (mañana o tarde). Al comienzo se informó a los estudiantes del objetivo del estudio y sus fines puramente académicos, se explicó lo que se pedía de ellos y se solicitó explícitamente su participación voluntaria. El tiempo típico empleado en la toma de datos en cada grupo de clase fue de 40 minutos, con un máximo de casi 50 minutos (evaluación de procesos de la ciencia).

Los datos recogidos fueron ordenados con Excel y tratados estadísticamente con el programa SPSS v.22. El estadístico ANOVA fue usado en los contrastes asociados con los factores y sus combinaciones.

En un análisis complementario se utilizó el estadístico t-Student de grupos emparejados en la comparación antes/después de la formación de las puntuaciones de cada bloque de contenidos y cada categoría procedimental.

El nivel crítico de la significación para el rechazo de la hipótesis nula fue siempre $p=,05$.

Resultados

La Tabla III proporciona los promedios de interés y las desviaciones típicas (entre paréntesis) para el conocimiento conceptual y para el procedimental (ambos con valores máximos 1).

TABLA III. Promedios (DT) para el conocimiento conceptual básico y para el conocimiento sobre procesos de la ciencia, en función de distintos factores considerados

Factores	Cto. Conceptual		Cto. Procedimental	
	Fase I	Fase 2	Fase I	Fase 2
Global	0,50 (0,14)	0,60 (0,12)	0,65 (0,13)	0,68 (0,12)
Bloque/Categoría				
I	0,50 (0,21)	0,63 (0,22)	0,76 (0,21)	0,82 (0,19)
II	0,54 (0,20)	0,62 (0,17)	0,52 (0,21)	0,54 (0,19)
III	0,45 (0,20)	0,55 (0,18)	0,59 (0,23)	0,61 (0,20)
IV	0,53 (0,23)	0,60 (0,21)	0,78 (0,18)	0,82 (0,17)
V	----	----	0,43 (0,26)	0,44 (0,25)
Esp. Pre-Grado				
Ciencias experimentales	0,58 (0,14)	0,67 (0,11)	0,69 (0,13)	0,72 (0,10)
Resto especialidades	0,49 (0,14)	0,58 (0,12)	0,65 (0,13)	0,67 (0,13)
Titul. Acceso				
Bachiller o Universidad	0,50 (0,14)	0,60 (0,13)	0,65 (0,13)	0,69 (0,12)
Otros	0,49 (0,14)	0,58 (0,12)	0,65 (0,14)	0,65 (0,13)
Carga curricular				
Uni. Valencia	0,53 (0,15)	0,64 (0,11)	0,66 (0,13)	0,71 (0,10)
Uni. Castellón y Alicante	0,48 (0,14)	0,57 (0,13)	0,65 (0,13)	0,67 (0,13)

Fuente: elaboración propia

Evolución del Conocimiento Conceptual básico de Ciencias

En la Tabla III anterior se observa que el incremento global de la puntuación de conocimiento conceptual pasó de 0,50 a 0,60 en la muestra sobre un máximo de 1 punto. Ello implica que se obtuvo un incremento relativo, sobre el máximo posible, dado por:

$$\frac{(0,60 - 0,50)}{(1 - 0,50)} = 0,20$$

Es decir, se incrementó la puntuación sólo un 20% de lo que podría (teóricamente e idealmente) haberse incrementado.

El análisis de la varianza de la puntuación global, tomando como factor entre sujetos la “Fase” de formación en ciencias (antes/después de la formación obligatoria en el Grado de Magisterio), mostró que el aumento durante el Grado fue significativo y con un tamaño medio-alto del efecto ($F(1, 557)=71,388$; $p<,001$; $\eta^2=,114$), lo cual añade significado al valor obtenido para el aumento relativo del 20% en la puntuación global.

Como información complementaria, la interacción “Fase X Bloque conceptual” no fue significativa ($F(3, 555)= 1,958$; $p>,10$), de modo que, con las reservas necesarias por la falta de fiabilidad en los promedios de cada bloque, el incremento de conocimiento fue estadísticamente similar en los diferentes bloques conceptuales.

Evolución del Conocimiento sobre Procesos elementales de la Ciencia

Las distintas habilidades procedimentales consideradas en el instrumento presentaron diferentes promedios antes y después de la formación en ciencias, como muestra la Tabla III. El cambio en la puntuación global observado en este estudio fue de sólo 0,03 puntos (con puntuaciones máximas de 1 punto). El cambio relativo puede cuantificarse como sigue:

$$\frac{(0,68 - 0,65)}{(1 - 0,65)} = 0,086$$

Es decir, el cambio relativo fue cercano al 9% del total que podría (teóricamente; idealmente) haberse obtenido. El ANOVA del promedio

global tomando como factor la “Fase” del estudio, mostró un cambio significativo antes y después de la formación en ciencias durante el Grado, debido a que las mejoras antes/después fueron sistemáticas en la muestra ($F(1, 592)=9,091$; $p=,03$; $\eta^2=,015$). Sin embargo, este efecto tuvo un tamaño del efecto muy pequeño, en coherencia con los valores pequeños de las mejoras observadas.

Un análisis multivariante complementario se realizó tomando como factor entre-sujetos la Fase y como factor intra-sujetos las cinco Categorías de procesos elementales de la ciencia. En este estudio de fiabilidad restringida, la interacción “Fase” X “Categoría” no fue significativa (Lambda Wilks: $F(4, 589)=2,021$; $p>,05$) lo que sugiere que todas las categorías procedimentales mostraron una mejora prácticamente equivalente en la comparación antes/después.

Efectos de la Especialidad de los estudios anteriores al Grado de Magisterio

Antes de iniciarse la formación en ciencias, el promedio global en conocimiento conceptual mostró una dependencia significativa de la especialidad o rama de los estudios pre-grado, comparando Ciencias experimentales con el Resto ($F(1,283)=18,960$; $p<,001$; $h^2=,063$) y con un tamaño del efecto medio-bajo.

Lo mismo sucedió con el conocimiento procedimental antes de la formación ($F(1, 306)= 4,343$; $p=,038$; $\eta^2=,014$).

Al acabar la formación obligatoria en ciencias en el Grado, las diferencias en ambos tipos de conocimiento se mantuvieron significativas, tanto para conocimiento conceptual ($F(1, 268)=22,950$; $p<,001$; $\eta^2=,079$) como procedimental ($F(1, 284)=9,073$; $p=,003$; $h^2=,031$).

La interacción entre la Especialidad de los estudios pre-grado y el factor “Fase” no fue significativa en ambos tipos de conocimiento ($F<1$), lo que indica que las diferencias observadas debidas a la especialidad de los estudios pre-Grado antes de la formación se mantuvieron tras la formación obligatoria en el Grado de Maestro de Primaria.

Efectos de la Titulación de acceso al Grado de Maestro de Primaria

Se realizaron sendos ANOVA para las puntuaciones en conocimiento conceptual y procedimental en relación con el factor Titulación de acceso al Grado. En ambos casos no hubo diferencias significativas en la fase anterior a la formación en el grado ($p > ,10$) ni tampoco interacción significativa ($p > ,10$) entre Fase y Titulación (Bachillerato & Universidad/ Otros). Globalmente, la mejora en conocimientos científicos de los alumnos procedentes de Bachillerato o de otro grado universitario, no difiere de la de aquellos que accedieron desde los módulos formativos de Grado Superior u otra vía.

Efectos debido a diferencias en la carga curricular de las universidades

Como se mencionó antes, hay diferencias en el número de horas destinado a la formación en ciencias, y también en el modo en que se distribuyen en asignaturas, entre la Universidad de Valencia y las otras dos, Castellón y Alicante. En ambas fases de la formación los promedios fueron superiores en la universidad de Valencia que en las otras universidades conjuntamente Castellón y Alicante. Sin embargo, la interacción entre la Fase y la distinta carga en los currículos no fue significativa, ni para conocimiento conceptual, ni para el procedimental ($p > ,10$ en ambos casos). Por tanto, el cambio antes/después de la formación fue similar con independencia de que la carga lectiva fuera mayor en Valencia y menor en Castellón y Alicante.

Discusión y Conclusiones

El objetivo del presente trabajo fue evaluar si los maestros en formación inicial en la Comunidad Valenciana están alfabetizados científicamente como para alfabetizar a su vez a quienes deben educar. Para ello, se atendió a los niveles de conocimiento en ciencia básica que presentan los y las estudiantes de la Comunidad Valenciana antes y después de su formación en el Grado de Maestro de Primaria. Los resultados en esta Comunidad no tienen por qué ser idénticos a los de otras, pero son una

buena representación que permite realizar una aproximación al nivel de alfabetización científica que presentan los maestros/as en formación en el Estado Español. El error muestral de nuestros análisis se sitúa en torno al 5%, muy similar al de estudio FECYT (2015) para la Comunidad Valenciana (5,03%). Se tomaron dos de los indicadores de alfabetización científica propuestos por Miller (1998), el conocimiento de conceptos científicos básicos y la comprensión de los procesos científicos, para los que se utilizaron instrumentos previamente validados (Verdugo-Perona, Solaz-Portolés y Sanjosé, 2016; Monde-Mónica, 2005).

Los resultados indican que los futuros/as maestros/as de educación primaria acceden al Grado con unos conocimientos conceptuales de nivel general bajo si se tiene en cuenta que la evaluación se realizó sobre contenidos de ciencia básica, típicos de los currículos de Primaria (no de Bachiller ni de Universidad). Los promedios aumentaron significativamente tras la formación en el Grado pero los tamaños del efecto fueron medio en el conocimiento conceptual, o muy pequeño en el conocimiento procedimental.

El presente trabajo no permite conocer las razones del pequeño tamaño de la mejora en el caso del conocimiento procedimental, pero uno de los factores que debería tenerse en cuenta para futuros estudios podría ser la tradicional escasa relación de los estudiantes con el trabajo indagatorio durante el Grado de Maestro de Primaria, como también sucede en otros contextos (Vilches y Gil-Pérez, 2007; Abd-El-Khalick et al, 2004).

El conocimiento conceptual mostrado por los participantes arrojó promedios bajos, si se tiene en cuenta que la evaluación se realizó sobre contenidos de ciencia básica, típicos de los currículos de Primaria. El nivel promedio aumentó significativamente tras la formación obligatoria en el Grado, con un tamaño del efecto medio.

Los niveles de conocimiento conceptual mostrados en el presente estudio son similares a los obtenidos en estudios generales anteriores (Eurobarómetro 224). Sin embargo, son claramente inferiores a los obtenidos en el estudio FECYT (2015). Las diferencias pueden radicar en el grado de dificultad de las preguntas realizadas en ambos estudios. Por un lado, FECYT utiliza aseveraciones constituidas por ideas completas que deben ser “reconocidas” como verdaderas o falsas por los participantes. Sin embargo, en nuestro estudio el participante debe discriminar un concepto entre otros tres distractores pertenecientes al mismo campo

conceptual, o a uno cercano. Por otro lado, el cuestionario empleado en este estudio (Verdugo-Perona, Solaz-Portolés y Sanjosé, 2016) implica mayor cantidad de conceptos que el empleado en FECYT. También es cierto que esa mayor dificultad es coherente con la formación que parece exigible a un futuro maestro de primaria. Los maestros en formación mostraron un nivel de dominio procedimental superior al de la población general objeto de estudio de FECYT. Este último estudio está basado en un sólo ítem fundamental y general (“entender la importancia de comparar grupos experimentales con grupos de control”). A diferencia de este estudio con un solo ítem, y de acuerdo con las características más específicas de nuestros participantes (universitarios, maestros en formación inicial), nuestro estudio utilizó un cuestionario validado con una multiplicidad y variedad de situaciones (Monde-Mónica, 2005), lo cual permite mayor profundidad y fiabilidad en la evaluación.

El análisis de los resultados por bloques o categorías de contenidos permiten identificar aquellos con peores resultados y que requieren de especial atención. En cuanto a contenidos procedimentales, la categoría V (“Diseño de experimentos”) no solo ha obtenido los peores resultados, sino que las competencias de los estudiantes en esta categoría prácticamente no mejoran durante el Grado. Este resultado puede explicarse, primero, por la dificultad en adquirir competencias de este cariz, propias de investigadores y no de maestros, y, segundo, quizás porque durante el Grado no se tienen muchas oportunidades de trabajar el diseño de experimentos, que es una tarea que demanda mucho tiempo. A pesar de que desde hace décadas se viene proponiendo el uso de metodologías basadas en la indagación para la enseñanza de las ciencias (Osborne y Dillon, 2008), aún no hay mucho tratamiento de las competencias necesarias para ello en maestros en formación (Tierno, Gavidia y Tuzón, 2018; Haefner y Zembal-Saul, 2004). Por supuesto, debemos pensar que los profesores universitarios atienden problemas urgentes con sus estudiantes, como la escasa preparación conceptual específica, que obstaculizan un aprendizaje didáctico que vaya más allá de lo superficial. En nuestro estudio, los participantes mostraron un conocimiento conceptual básico en ciencias claramente mejorable, y con resultados especialmente bajos en el bloque III (“Salud e higiene personal”), al que tradicionalmente se destinan una gran cantidad de recursos que, al parecer, no están dando los resultados deseados. Se realizaron análisis complementarios de algunos factores que podrían

matizar los resultados generales obtenidos. El tipo de titulación de acceso al Grado de Magisterio (Bachillerato y otras Licenciaturas, frente al resto de opciones Módulos Formativos, Acceso Mayores, etc.) no produjo discriminación alguna en los resultados obtenidos, ni en conocimiento conceptual ni en procedimental. Esto indica que tanto quienes acceden desde los Módulos Formativos como quienes lo hacen por acceso de mayores, muestran niveles de conocimiento similares al alumnado que procede del Bachillerato.

Las diferencias estriban en la especialidad de los estudios, y no tanto en el tipo de estudios. En efecto, los alumnos con una formación pregrado especializada en ciencias experimentales iniciaron los estudios de Magisterio con un nivel de conocimientos significativamente mejor que el resto, tanto en conceptos como en procesos de la ciencia. Esos estudiantes se beneficiaron más que el resto de la enseñanza recibida durante el Grado y continuaron manteniendo un nivel de conocimientos científicos más elevado al finalizar la formación. Este efecto beneficioso del conocimiento previo es bien conocido y ha sido replicado en numerosos estudios educativos (Braasch y Goldman, 2010; Irrazabal, 2010) Por lo visto es necesario prestar una especial atención en el Grado a las personas con menor conocimiento previo en ciencias, y tratar de mejorar su motivación, esfuerzo, atención y goce hacia las ciencias.

Finalmente, en la universidad que oferta mayor número de créditos en la formación obligatoria de ciencias experimentales (la de Valencia), se mostró mayor nivel de dominio conceptual y procedimental que en las otras universidades estudiadas, tanto antes como tras la formación en el Grado. El mayor número de créditos no supuso un aumento mayor del conocimiento respecto de las otras dos universidades, en especial comparando con la de Alicante. Obviamente, el número de créditos ofertados y cursados durante el Grado no parece un factor suficiente para explicar el aprendizaje de los estudiantes. La metodología del profesorado, el modo en que gestionan sus recursos y el mayor o menor énfasis puesto en determinados contenidos (conceptuales, procedimentales o didácticos) son, seguramente, factores clave que no han sido contemplados en este estudio evaluativo.

Un factor que podría colaborar a una mayor eficiencia instruccional en Alicante respecto de Valencia es el hecho de que en aquella facultad los conocimientos disciplinares se imparten conjuntamente con los conocimientos didácticos. Sin duda “ponerse en el lugar del profesor

para explicar” provoca en los estudiantes un fenómeno metacognitivo, lo que mejora el conocimiento disciplinar. Este fenómeno ha sido puesto de relieve en varios estudios (Dignath y Büttner, 2008).

Limitaciones del estudio

El presente estudio, aunque de suficiente validez externa y fiabilidad en cuanto a los resultados generales, presenta ciertas características que definen sus limitaciones.

En primer lugar, en este trabajo se obtuvo una evaluación independiente de dos componentes del conocimiento científico: el conceptual y el procedimental. Aunque ello resulta técnicamente importante para poder concluir acerca de dichos componentes por separado, es cierto que el conocimiento científico y su utilización en la vida diaria se basan en la conjunción de conceptos y procesos para describir, comprender o predecir fenómenos naturales. El hecho de que el instrumento empleado para evaluar el conocimiento conceptual haga intervenir la memoria del sujeto, puede influir en los bajos promedios obtenidos, incluso al final de la formación en el Grado. De modo similar, el instrumento utilizado para evaluar el conocimiento sobre procesos de la ciencia no involucra conocimiento conceptual específico para contestar correctamente cada ítem, por lo que los resultados podrían variar si las situaciones presentadas a los participantes implicasen también conceptos científicos.

En segundo lugar, los instrumentos empleados sólo garantizan fiabilidad suficiente en los promedios generales, pero no garantizan fiabilidad en las puntuaciones referidas a cada bloque conceptual o a cada categoría procedimental de la ciencia. Por tanto, y como se ha explicitado en su momento, las evaluaciones que se pretenda realizar de cada uno de esos bloques o categorías deben tomarse como meras aproximaciones u orientaciones. Hay que esperar a obtener suficientes replicaciones para aumentar esa fiabilidad, o utilizar instrumentos más específicos (por ejemplo, con mayor número de ítems bien cohesionados) para cada bloque conceptual o cada categoría procedimental.

En tercer lugar, queda pendiente la evaluación y análisis en esta misma población del conocimiento sobre Naturaleza de la Ciencia, que ha merecido mucha atención (Fernández, Gil, Carrascosa y Cachapuz, 2002; Irez, 2006; Guisasola y Morentin, 2007; Acevedo-Díaz, 2008).

También sería de interés evaluar el conocimiento didáctico en ciencias experimentales (Blanco, Mellado y Ruiz, 1995; Loughran, Mulhall y Berry, 2008; Nilsson, 2008) en maestros en formación. Ambos tipos de conocimiento son importantes para la futura evolución de la cultura científica de la población en general.

La última limitación destacable es la elección de un estudio sincrónico, en lugar de diacrónico que disminuiría la varianza de error al emplear los mismos informantes en todo el proceso.

Referencias

- Abd-El-Khalick, F., Boujaoude, S., Duschl, R., Lederman, N. G., Mamlok-Naaman, R., Hofstein, A. y Tuan, H. L. (2004). Inquiry in Science Education: International perspectives. *Science Education*, 88(3), 397-419. DOI:10.1002/sce.10118
- Acevedo, J.A. (2004). Reflexiones sobre las finalidades de la enseñanza de las ciencias: Educación científica para la ciudadanía. *Revista Eureka sobre Enseñanza y Divulgación de las Ciencias*, 1, 3-16.
- Acevedo-Díaz, J. A. (2008). El estado actual de la naturaleza de la ciencia en la didáctica de las ciencias. *Revista Eureka sobre Enseñanza y Divulgación de las Ciencias*, 5(2), 133-169.
- American Association for the Advancement of Science [AAAS] (1989). *Science for all Americans*. Recuperado de: <http://www.project2061.org/publications/sfaa/online/sfaatoc.html>.
- Appleton, K. (2003). How do beginning primary school teachers cope with science? Toward an understanding of science teaching practice. *Research in science education*, 33(1), 1-25.
- Blanco, L. J., Mellado, V. y Ruiz, C. (1995). Conocimiento didáctico del contenido en Ciencias Experimentales y Matemáticas y Formación de Profesores. *Revista de Educación*, 307, 427-446.
- Braasch, J. L. y Goldman, S. R. (2010). The role of prior knowledge in learning from analogies in science texts. *Discourse Processes*, 47(6), 447-479. DOI:10.1080/01638530903420960

- Bransford, J., Brown, A. y Cocking, R. (Eds.). (1999). How people learn. Washington, DC: National Academy Press.
- Brossard, D. y Shanahan, J. (2006). Do they know what they read? *Science Communication*, 28(1), 47-63. DOI:10.1177/1075547006291345
- Dignath, C. y Büttner, G. (2008). Components of fostering self-regulated learning among students. A meta-analysis on intervention studies at primary and secondary school level. *Metacognition and Learning*, 3(3), 231-264. DOI:10.1007/s11409-008-9029-x
- Cantón, I., Cañón, R. y Arías, A. R. (2013). La formación universitaria de los maestros de Educación Primaria. *Revista Interuniversitaria de Formación del Profesorado*, 27(1), 45-63.
- Cañal, P. (2000). El conocimiento profesional sobre las ciencias y la alfabetización científica en Primaria. *Alambique Didáctica de las Ciencias Experimentales*, 24, 46-56.
- Cañal, P. (2008). ¿Cómo orientar la formación inicial del profesorado de primaria en didáctica de las ciencias experimentales? En M. R. Jiménez Liso (Ed.). *Ciencias para el mundo contemporáneo y formación del profesorado en Didáctica de las Ciencias Experimentales* (pp. 256-263). Almería: Ed. Univ. Almería.
- Criado, A. M. y García-Carmona, A. (2011). Las experiencias prácticas para el conocimiento del medio (natural y tecnológico) en la formación inicial de maestros. *Investigación en la Escuela*, 74, 73-88.
- Decreto 108/2014, de 4 de julio, del Consell, por el que establece el currículo y desarrolla la ordenación general de la educación primaria en la Comunitat Valenciana. [2014/6347]. Publicado en DOGCV, num 7311/07 07 2014, p 16510. Recuperado de: http://www.docv.gva.es/datos/2014/07/07/pdf/2014_6347.pdf.
- FECYT (2015). Percepción Social de la Ciencia y la Tecnología en España. Madrid. Recuperado de: <http://www.fecyt.es/es/publicacion/percepcion-social-de-la-ciencia-y-la-tecnologia-2014>.
- Fernández, I., Gil, D., Carrascosa, J. y Cachapuz, A. (2002). Visiones deformadas de la ciencia transmitida por la enseñanza. *Enseñanza de las ciencias*, 20(3), 477-488.
- García Barros, S., Bugallo, Á. y Fuentes, M. J. (2013). Los objetivos de las ciencias en primaria y las necesidades formativas, vistas por los maestros en formación. *Enseñanza de las Ciencias*, número extra, 1460-1466.

- García Barros, S., Martínez Losada, C., Vega, P. y Mondelo, M. (2000). Propuesta de intervención para la formación inicial del profesorado de Educación Primaria en Ciencias Experimentales. *Revista Interuniversitaria de Formación del Profesorado*, 38, 153-165.
- García-Ruiz, M. y Sánchez, B. (2006). Las actitudes relacionadas con las Ciencias Naturales y sus repercusiones en la práctica docente de profesores de primaria. *Perfiles Educativos*, 114, 61-89.
- Gil, D. y Vilches, A. (2006). Educación ciudadana y alfabetización científica: mitos y realidades. *Revista Iberoamericana de Educación*, 42, 31-53.
- Glaser, R. (1995). Application and theory: Learning theory and the design of learning environments. Paper presented at the 23rd International Congress of Applied Psychology, July 17-22, 1994, Madrid, Spain.
- Gómez, P. y Gutiérrez-Gutiérrez, A. (2014). Conocimiento matemático y conocimiento didáctico del futuro profesor español de primaria. Resultados del estudio TEDS-M. En M. T. González, M. Codes, D. Arnau y T. Ortega (Eds.), *Investigación en Educación Matemática XVIII* (pp. 99-114). Salamanca: SEIEM.
- Guisasola, J. y Morentin, M. (2007). ¿Comprenden la naturaleza de la ciencia los futuros maestros y maestras de Educación Primaria? *Revista Electrónica de Enseñanza de las Ciencias*, 6(2), 246-262.
- Haefner, L. A., y Zembal-Saul, C. (2004). Learning by doing? Prospective elementary teachers' developing understandings of scientific inquiry and science teaching and learning. *International Journal of Science Education*, 26(13), 1653-1674. DOI:10.1080/0950069042000230709
- Harlen, W. y Holroyd, C. (1997). Primary teachers' understanding of concepts of science: impact on confidence and teaching. *International Journal of Science Education*, 19(1), 93-105. DOI:10.1080/0950069970190107
- Hurd, P. D. (1998). Scientific Literacy: New minds for a changing world. *Science Education*, 82(3), 407-416. DOI:10.1002/(SICI)1098-237X(199806)82:3<407::AID-SCE6>3.0.CO;2-G
- Irez, S. (2006). Are we prepared? An assessment of preservice science teacher educators' beliefs about nature of science. *Science Education*, 90(6), 1113-1143. DOI:10.1002/sce.20156
- Irrazabal, N. (2010). La comprensión de textos expositivos en estudiantes universitarios: la función del conocimiento previo. *Revista de Psicología*, 12, 7-21.

- Jarvis, T. y Pell, A. (2004). Primary teachers' changing attitudes and cognition during a two-year science in-service programme and their effect on pupils. *International Journal of Science Education*, 26 (14), 1787-1811. DOI:10.1080/0950069042000243763
- Jurišević, M., Glažar, S. A., Pučko, C. R. y Devetak, I. (2008). Intrinsic Motivation of Pre-service Primary School Teachers for Learning Chemistry in Relation to their Academic Achievement. *International Journal of Science Education*, 30(1), 87-107.
- Kang, NH. (2007). Elementary teachers' teaching for conceptual understanding: learning from action research. *Journal of Science Teacher Education*, 18(4), 469-495. DOI: 10.1007/s10972-007-9050-y.
- Krugar, C., Summers, M. y Palacio, D. (1990) A survey of primary school teachers' conceptions of force and motion. *Educational Research*, 32(2), 83-95. DOI:10.1080/0013188900320201
- Laugksch, R.C. y Spargo, P. (1996). Construction of a paper-and-pencil test of basic scientific literacy goals recommended by the American Association for the Advancement of Science. *Public Understanding of Science*, 5(4), 331-359. DOI:10.1088/0963-6625/5/4/003
- Lloyd, J. K., Smith, R. G., Fay, C. L., Khang, G. N., Wah, L. L. K. y Sai, C. L. (1998). Subject knowledge for science teaching at primary level: a comparison of pre-service teachers in England and Singapore. *International Journal of Science Education*, 20(5), 521-532. DOI:10.1080/0950069980200502
- Loughran, J., Mulhall, P., y Berry, A. (2008). Exploring pedagogical content knowledge in science teacher education. *International Journal of Science Education*, 30(10), 1301-1320. DOI:10.1080/09500690802187009
- Martínez-Chico, M., Jiménez Liso, M. R. y 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. DOI:10.25267/Rev_Eureka_ensen_divulg_cienc.2015.v12.i1.10
- Mellado Jiménez, V. (1996). Concepciones y prácticas de aula de profesores de ciencias, en formación inicial de primaria y secundaria. *Enseñanza de las Ciencias*, 14(3), 289-302.
- Miller, J. D. (1989). Scientific literacy. Paper presented at the annual meeting of the American Association for the Advancement of Science, San Francisco. January 1989.

- Miller, J. D. (1998). The measurement of civic scientific literacy. *Public understanding of science*, 7(3), 203-223. DOI:10.1088/0963-6625/7/3/001
- Monde-Monica, K.M. (2005). Development and validation of a test of integrated science process skills for the further education and training learners. Tesis de Master. Universidad de Pretoria, Sudáfrica. Recuperado de: <http://repository.up.ac.za/bitstream/handle/2263/24239/dissertation.pdf?sequence=1>
- National Research Council. (1996). National Science Education Standards. Washington, DC: National Academy Press. Recuperado de: <http://www.nap.edu/catalog/4962/national-science-education-standards>.
- Murphy, C. y Smith, G. (2012). The impact of a curriculum course on pre-service primary teachers' science content knowledge and attitudes towards teaching science. *Irish Educational Studies*, 31(1), 77-95. DOI:10.1080/03323315.2011.634061
- National Research Council. (1996). National science education standards. National Academies Press.
- National Science Board (2014). Science and Technology: Public attitudes and understanding. En *Science and Engineering Indicators 2014* (7.20-7.23). Arlington VA: National Science Foundation (NSB-14-01). Recuperado de: <http://nsf.gov/statistics/seind14/index.cfm/chapter-7/c7h.htm>.
- Nilsson, P. (2008). Teaching for Understanding: The complex nature of pedagogical content knowledge in preservice education. *International Journal of Science Education*, 30(10), 1281-1299. DOI:10.1080/09500690802186993
- Osborne, J. y Dillon, J. (2008). *Science education in Europe: Critical reflections* (Vol. 13). London: The Nuffield Foundation
- Papageorgiou, G., Stamovlasis, D. y Johnson, P. M. (2010). Primary Teachers' Particle Ideas and Explanations of Physical Phenomena: Effect of an in-service training course. *International Journal of Science Education*, 32(5), 629-652. DOI:10.1080/09500690902738016
- Pérez de Landazábal, M.C., Benegas, J., Cabrera, J.S., Espejo, R., Macías, A., Otero, J., Seballos, S. y Zavala, G. (2010). Comprensión de conceptos básicos de la Física por alumnos que acceden a la universidad en España e Iberoamérica: limitaciones y propuestas de mejora *Latin American Journal of Physics Education*, 4 (3), 655-668. Recuperado de: <http://www.lajpe.org>.

- Porlán, R.; Martín del Pozo, R.; Rivero, A.; Harres, J.; Azcárate, P. y Pizzato, M. (2010). El cambio del profesorado de ciencias I: Marco teórico y formativo. *Enseñanza de las Ciencias*, 28(1), 31-46.
- Rivero, A., Hamed, S., Martín del Pozo R., Solís, E. Fernández, J., Porlán R., Rodríguez, F., Solís, C., Azcárate, P. y Ezquerro, Á. (2013). La formación inicial de maestros de primaria: qué hacer y cómo en didáctica de las ciencias. *Enseñanza de las Ciencias, número extra*, 3045-3050.
- Rundgren, C.-J., Rundgren, S.-N. C., Tseng, Y.-H., Lin, P.-L. y Chang, C.-Y. (2010). Are you SLiM? Developing an instrument for civic scientific literacy measurement (SLiM) based on media coverage. *Public Understanding of Science*, 20(10), 1-15. DOI:10.1177/0963662510377562. Recuperado de: https://www.researchgate.net/profile/Chun_Yen_Chang/publication/247156553_Are_you_SLiM_Developing_an_instrument_for_civic_scientific_literacy_measurement_SLiM_based_on_media_coverage/links/545972bd0cf2bccc4912bcee/Are-you-SLiM-Developing-an-instrument-for-civic-scientific-literacy-measurement-SLiM-based-on-media-coverage.pdf
- Sáenz Castro, C. (2007). La competencia matemática (en el sentido de PISA) de los futuros maestros. *Enseñanza de las Ciencias*, 25(3), 355-366.
- Special Eurobarometer 43 (1990). Les Européens, la Science et la Technologie. Commission of the European Communities. Paris. Recuperado de: <http://ec.europa.eu/COMMFrontOffice/PublicOpinion/index.cfm/Survey/getSurveyDetail/yearFrom/1988/yearTo/1990/surveyKy/101>].
- Special Eurobarometer 224 (2005). Europeans, Science and Technology. Commission of the European Communities. Brussels. Recuperado de: <http://ec.europa.eu/COMMFrontOffice/PublicOpinion/index.cfm/Survey/getSurveyDetail/yearFrom/2004/yearTo/2007/surveyKy/447/p/3>].
- Special Eurobarometer 76 (1993). Europeans, Science and Technology. Commission of the European Communities. Brussels. Recuperado de: <http://ec.europa.eu/COMMFrontOffice/PublicOpinion/index.cfm/Survey/getSurveyDetail/yearFrom/1991/yearTo/1993/surveyKy/131>]
- Summers, M. y Krugar, C. (1992). Research into English primary teachers' understanding of the concept of energy. *Evaluation and Research in Education*, 6, 95-111. DOI:10.1080/09500799209533321

- Tierno, S., Gavidia, V. y Tuzón, P. (2018). Radiografía de la enseñanza de las ciencias en las Facultades de Magisterio de España. Manuscrito en preparación.
- Verdugo-Perona, J.J., Solaz-Portolés, J.J. y Sanjosé, V. (2016). Pre-Service primary school teachers' science content knowledge: an instrument for its assessment. *International Journal of Innovation in Science and Mathematics Education*, 24(2), 37-51.
- Vilches, A. y Gil-Pérez, D. (2007). La necesaria renovación de la formación del profesorado para una educación científica de calidad. *Tecné, Episteme y Didaxis*, 22, 67-85.

Dirección de contacto: José Javier Verdugo Perona. Florida Universitaria, Unidad de Educación. Carrer del Rei en Jaume I, 2, 46470 Catarroja, València. E-mail: javiverpe@gmail.com

Anexo

Instrumento usado para evaluar el conocimiento conceptual básico en ciencias

Los contenidos considerados son: Traslación; Galaxia; Fotosíntesis; Alveolos pulmonares; Uretra; Refracción; Ciclo lunar; Chimenea volcánica; Sépalos; Riñones; Sistema nervioso central; Energía renovable; Clima; Animales invertebrados; Vasos liberianos; Glóbulos rojos; Hígado; Cambios químicos; Mineral; Órganos; Virus; Cigoto; Glúcidos; Tono; Anemómetro; Función biológica de las flores; Células vegetales y animales; Intestino delgado; Filtración; Mezcla de colores

Algunos ítems:

PREGUNTAS	OPCIONES DE RESPUESTA
¿Qué nombre recibe el movimiento que la Tierra realiza alrededor del Sol?	a) Rotación b) Precesión c) Traslación* d) Circunferencia
¿Qué parte de un volcán es la chimenea volcánica?	a) Cámara subterránea en la que se acumula el magma b) Orificio de salida al exterior de un volcán c) Conducto por el que asciende el magma desde el interior terrestre* d) Conjunto de rocas y cenizas que se forma al enfriarse la lava
¿Qué nombre reciben las hojas que forman el cáliz de una flor?	a) Sépalos* b) Estambres c) Pétalos d) Corola
¿Qué componente de la sangre se encarga de transportar el oxígeno?	a) Plasma b) Plaquetas c) Glóbulos blancos d) Glóbulos rojos*
¿Cómo se conoce al cambio de dirección que se produce en un rayo luminoso al pasar de un medio material a otro?	a) Reflexión b) Difracción c) Atenuación d) Refracción*

¿Qué clase de cambios causan las modificaciones en la composición de la materia?

- a) Los cambios físicos
- b) Los cambios químicos*
- c) Los cambios biológicos
- d) Ninguno de ellos, la materia no varía

Instrumento empleado para evaluar el conocimiento sobre procesos de la ciencia

El instrumento completo en inglés puede ser obtenido en la página web: <http://repository.up.ac.za/bitstream/handle/2263/24239/dissertation.pdf?sequence=1>

Algunos ítems:

PREGUNTA	OPCIONES DE RESPUESTA
En un anuncio de radio se afirma que Surf (un detergente) produce más espuma que otros tipos de jabón en polvo. Raquel quería confirmar esta afirmación. Puso la misma cantidad de agua en cuatro cuencos y añadió 1 taza de un tipo diferente de jabón en polvo (incluyendo Surf) a cada cuenco. Agitó vigorosamente el agua en cada cuenco y observó el que más espuma producía. ¿Cuál de los factores de abajo probablemente NO afecta a la producción de espuma del jabón en polvo?	<ul style="list-style-type: none">a) La cantidad de tiempo usado para agitar el aguab) La cantidad de agitación realizadac) El tipo de cuenco utilizado*d) El tipo de jabón en polvo utilizado
Un jardinero escolar corta el césped de 7 campos de fútbol diferentes. Cada semana, corta un campo diferente. El césped es habitualmente más largo en algunos campos que en otros. Él hace algunas suposiciones sobre por qué la altura del césped es diferente. ¿Cuál de las siguientes es una explicación comprobable adecuada para la diferencia en la altura del césped?	<ul style="list-style-type: none">a) Los campos que reciben más agua tienen el césped más largo*b) Los campos que tienen el césped más corto son más adecuados para jugar al fútbolc) Cuantas más piedras hay en el campo, más difícil es cortar el céspedd) Los campos que absorben más dióxido de carbono tienen el césped más largo

Un estudiante quería saber si un incremento en la cantidad de vitaminas dada a los niños conlleva mayor crecimiento. ¿Cómo puede calcular el estudiante la rapidez con la que crecerán los niños?

Una profesora de biología quería mostrar a su clase la relación entre la intensidad luminosa y la tasa de crecimiento de las plantas. Llevó a cabo una investigación y obtuvo los siguientes resultados.

Intensidad de la luz (Candela)	Tasa de crecimiento de la planta (cm)
250	2
800	5
1000	9
1200	11
1800	12
2000	15
2400	13
2800	10
3100	5

Tabla 1.5. Muestra la relación entre la intensidad de la luz y la tasa de crecimiento de una planta.

¿Cuál de los siguientes enunciados describe correctamente lo que muestran los resultados?

- a) Contando el número de palabras que los niños pueden decir a una determinada edad
 - b) Pesando la cantidad de vitaminas dada a los niños
 - c) Midiendo los movimientos de los niños
 - d) Pesando los niños cada semana*
- a) A medida que la intensidad luminosa aumenta, el crecimiento de la planta también aumenta
 - b) A medida que la planta crece, la intensidad luminosa disminuye.
 - c) A medida que la planta crece, la intensidad luminosa aumenta y después disminuye
 - d) A medida que la intensidad luminosa aumenta, el crecimiento de las plantas aumenta y después disminuye*



Reseñas

Martínez, M., Esteban, F. Jover, G. y Payá, M. (2016). *La Educación, en teoría*. Madrid: Síntesis. 219 pp. ISBN: 978-84-9077-438-0.

Situar a la teoría en supuestos controvertidos derivados de la práctica y pensar, pensar sobre la educación, “una y otra vez, sin descanso” (p. 11). Este es el propósito contenido de esta obra orquestada por cuatro autores referentes en la disciplina de la Teoría de la Educación. El libro que se reseña nos introduce en la reflexión educativa que no siempre se presenta como inspira la teoría y que, sin embargo, no puede encontrarse más necesitada de la misma.

La obra posa su mirada en cuatro grandes pilares temáticos constituidos por: el alumnado, los docentes, lo que se enseña y se aprende y finalmente, la práctica educativa. El primero de los bloques temáticos, *¿Para qué los profesores?*, aterriza con una reverencia a la figura del docente y su capacidad de influencia personal: *Sin usted nada de esto hubiera sucedido*. Se invita a pensar la docencia como una forma de vida que se caracteriza por su excelencia personal a la hora de: a) transmitir lo mejor, b) crear espacios únicos de aprendizaje y c) dispensar un trato personal con el alumnado. Los autores se aproximan a la realidad educativa bajo la premisa del predominio latente de la racionalidad instrumental en las instituciones educativas y, como consecuencia, esto deriva en la pérdida de la dimensión personal y la identidad del alumnado. La reflexión sobre la relación educativa desvela un vínculo fundamentado en la ayuda, la amistad, el respeto, la confianza, el optimismo, la igualdad, la escucha y la reciprocidad, orientados hacia una finalidad inmanente que apunta a un objetivo a conseguir en la búsqueda de una influencia educativa.

En el segundo bloque se produce un giro hacia la cuestión de *¿Qué es necesario aprender hoy?* Como consecuencia del principio de la racionalidad instrumental que se configura bajo las coordenadas de eficacia y eficiencia, todo lo que es contemplado como útil es alineado como verdad. Los autores alertan sobre el riesgo de eclipsar en el conocimiento

educativo todo aquello relacionado con el pensamiento y su ejercicio de reflexión algo que, en el fondo, supone un desprecio a la cultura clásica. No obstante, ¿acaso todo lo verdadero podría ser considerado como útil? Es precisamente la experiencia con la cultura clásica la que permite establecer una continuidad universal entre distintas generaciones. Si la educación tiene el deber de proliferar las posibilidades para una vida más plena, ¿por qué reducir nuestra acción al territorio de lo útil? No se trata de lo útil, sino de lo valioso, todo conocimiento escolar debería ser valioso para el alumnado

La tercera parte asoma la mirada hacia la persona que aprende, el alumnado. Dos son las tradiciones que han prevalecido en la manera de pensar a los estudiantes: liberal y comunitarista. Con cierto predominio de la perspectiva liberal, advierten sobre los efectos negativos de este desequilibrio para la educación, especialmente para su dimensión moral. Si bien ambas han convivido de manera descompensada en las prácticas educativas, los autores desvelan su naturaleza complementaria y compatible. En su análisis no falta la confrontación entre los dos modos de entender la educación: como proceso de civilización y como desarrollo natural, ambos con implicaciones educativas muy distintas.

Una vez esbozada la reflexión sobre la figura del docente, del discente y de lo que se debería enseñar y aprender, el bloque final hace referencia a las prácticas educativas. Se establece una crítica a en torno a tres ejes principales que gobiernan en el panorama educativo: a) la ofuscación por la adecuación al entorno, b) la afirmación de que la práctica educativa es de todos y, por último, c) el excesivo protagonismo del alumnado en su educación. Finalmente, al último capítulo le sigue un breve peregrinaje histórico que capta la esencia sobre la configuración de la Teoría de la Educación como campo disciplinar y curricular en España. En su conjunto, la suma de los capítulos hace una lectura interactiva, abierta y flexible a todo tipo de públicos especialmente indicada para esas “mentes intranquilas y hambreadas” (p.10) que, sedientas por su inquietud e insatisfacción, alientan y enriquecen los coloquios que sostienen los derroteros educativos.

Laura Camas Garrido

Prendes Espinosa, M.P. & Román García, M. (Coords.) (2017). *Entornos personales de aprendizaje: Una visión actual de cómo aprender con tecnologías*. Barcelona: Octaedro. 181 pp. ISBN: 978-84-9921-901-1

El concepto de entorno personal de aprendizaje (PLE, en inglés) va más allá de una simple recopilación de herramientas digitales o una plataforma donde representar los recursos usados en los procesos de aprendizaje, debemos entenderlo considerando los procesos cognitivos, las estrategias y actitudes personales. Son una representación personal de todo aquello que afecta a nuestro aprendizaje dentro o fuera del aula, de forma presencial o virtual.

La obra cuenta con la colaboración de un total de 21 autores y se estructura en dos partes. La primera de ellas nos sitúa en CAPPLE, un proyecto financiado por el Ministerio de Economía y Competitividad (EDU2012-33256) donde se analizan las competencias para el aprendizaje permanente basado en el uso del PLE en los alumnos de último curso de grado de las universidades españolas. La segunda parte ofrece una visión ampliada mediante las aportaciones de autores externos, con otras visiones o concreciones sobre los PLE y estableciendo relaciones con otros proyectos o entornos.

El primer bloque, *El proyecto CAPPLE*, está conformado por cuatro capítulos que gradualmente nos permiten profundizar en la comprensión de los PLE y del proyecto CAPPLE. En el primer capítulo, las coordinadoras del libro ofrecen una aproximación general al concepto y nos introducen al proyecto desarrollado. El segundo capítulo, de la mano de González, Sánchez y Castañeda, complementa el anterior y lo amplía con una revisión de teorías y caracterización de las investigaciones que se han hecho previamente sobre los PLE. En tercer lugar, encontramos un espacio mayormente dedicado a explicar la creación y validación del cuestionario utilizado para llevar a cabo la investigación. Por último, en el cuarto capítulo, Gutiérrez, López, Serrano y Solano presentan resultados destacados de los datos obtenidos así como una serie de propuestas fundamentadas dirigidas a las instituciones universitarias para la mejora de los PLE.

El segundo bloque se titula *Aportaciones y visiones en torno al PLE* y nos permite profundizar en el concepto a lo largo de cuatro capítulos más, con aportaciones de autores destacados en el ámbito. Salinas y Marín nos

presentan los PLE mediante una revisión de los LMS, remarcando que la evolución pone el foco en que los alumnos ya no se deben adaptar a un sistema, sino que cada alumno configura su propio espacio. En el sexto capítulo, Aguiar, Artiles, y Rodríguez reflexionan sobre enfoques didácticos y organizativos desde los entornos de aprendizaje en la formación del profesorado de enseñanza superior. Llegamos al séptimo, donde Galván y Molas centran su aportación en poner de relieve la relación entre PLE y portafolios digital; lo hacen ayudados de la ejemplificación mediante el proyecto de Carpeta Digital.

Finalmente, el octavo capítulo del monográfico, que se enmarca dentro de la segunda parte del libro, tiene un formato diferenciado dado que se ofrecen *Visiones en torno al PLE* de forma individual por parte de autores internacionales. Bartolomé focaliza su aportación en la calidad de autorregulación del aprendizaje que indiscutiblemente a lo largo del libro comprobamos como elemento esencial de los PLE. Ebner ofrece su aportación con el conocimiento de su participación en los proyectos PLE de la Universidad Tecnológica de Graz. Martínez aporta la visión más crítica del libro poniendo a debate reflexiones que no deberían obviarse. Cabreiro versa sobre los PLE “en el camino hacia aprender a aprender”. Y, por último, Attwell pone la mirada en el futuro remarcando las claves en la relación del PLE con los MOOC, la educación abierta y el *learning analytics*, entre otros.

En suma, se trata de un monográfico que puede servir a distintos perfiles de lector que quieran profundizar en el conocimiento de los entornos personales de aprendizaje. Los estudiantes o perfiles noveles en el campo encontrarán múltiples visiones y conceptualizaciones para facilitar la comprensión de los PLE, mientras que investigadores, profesores o perfiles que tengan buen conocimiento del tema podrán profundizar en una conceptualización completa a la par que descubrir datos y conclusiones extraídas del proyecto nacional CAPPLE. La variedad de autores y la composición del monográfico lo convierten en un ejemplar de interés general.

Jordi Mogas Recalde

La *Revista de Educación* es una publicación científica del Ministerio de Educación y Formación Profesional 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 Atención al Ciudadano, 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.educacionyfp.gob.es/revista-de-educacion